ON DERIVING THE LEXICON

by

Richard William Sproat

B.A. University of California, San Diego

(1981)

Submitted to the Department of
Linguistics and Philosophy
in Partial Fulfillment of the
Requirements of the
Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1985

© Richard William Sproat 1985

The author hereby grants to M.I.T. permission to reproduce and to distribute copies of this thesis document in whole or in part.

Signature of Author Richard W. Sproat
Department of Linguistics and Philosophy
August 9, 1985

Certified by Kenneth Hale
Thesis Supervisor

Accepted by Morris Halle
Chairman, Departmental Graduate Committee

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

OCT 31 1985

LIBRARIES

Archives

ON DERIVING THE LEXICON

by

RICHARD WILLIAM SPROAT

Submitted to the Department of Linguistics and Philosophy
on August 9, 1985 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Linguistics

ABSTRACT

In this thesis I argue against the view that there is a separate word-formation component of the grammar, a component which has usually been termed the "Lexicon" in the recent morphological literature. I argue rather that the lexicon is what has been called the permanent lexicon, namely the data structure containing the information about stems and affixes and idiosyncratic compositions of the language, and that word formation is actually split between the syntax and the phonology in that it is principles of syntax which determine the syntactic well-formedness of words, and principles of phonology which determine phonological well-formedness.

In the first chapter I argue that morphological Bracketing Paradoxes, including phonological cliticization provide crucial evidence that words must have at least two levels of representation, in particular a syntactic one and a phonological one. I show that a simple Mapping Principle governs the relationship between the two levels of representation. Subsequently, in Chapters 2 and 3 I show that the syntactic well-formedness and behavior of a word, like that of a phrase, can be determined by syntactic principles including, but not limited to, X-bar Theory, Theta-Theory, Binding Theory, Case Theory and the Projection Principle.

In Chapter 4 it is argued that lexical phonology is not in principle distinct from what has come to be termed post-lexical phonology, in that principles such as Cyclicity and Bracketing Erasure, which have generally been argued to be hallmarks of lexical phonology are either not needed or not unique to lexical phonology. I argue too that the theory of Lexical Phonology cannot be taken to be a theory of word formation, but at most is a theory of phonological well-formedness. I propose an alternative to Lexical Phonology based upon the distinction between phonological words and stems

Finally, in the fifth chapter I summarize the psychological evidence pertaining to word-formation. I argue that the approach to morphology

taken here is at least as compatible with such evidence as other theories of word-formation, such as Lexical Phonology. I also discuss some residual conceptual issues raised by the approach taken in this thesis.

Thesis Supervisor: Kenneth Locke Hale

Title: Ferrari P. Ward Professor of Linguistics

Acknowledgments.

I should be very surprised indeed if I turned out to be the only person who found the Acknowledgments section of their thesis to be one of the most difficult parts to write: the greatest problem, needless to say, is that there is always the chance that someone's name is going to end up being omitted from the list, someone, perhaps, who in a seemingly insignificant way contributed greatly to the amelioration of being a graduate student in general, or to the production of the thesis in particular, the latter being in many ways the culmination of the former. Still, no guts no glory: I trust it is not too cliched to express the hope that nobody got left out who should be here.

I wish first to acknowledge the people who have aided me most in the production of this dissertation. Ken Hale, my thesis supervisor was a constant source of ideas; he read carefully all that I wrote and commented on everything. Particularly useful has been his unequaled knowledge of languages; he often commented on the implications of my ideas for constructions in languages I had never even thought of looking at, and the fact that he has never produced any strong counterevidence from this store of knowledge has given me some amount of confidence that my ideas must have somewhat on the right track. My other committee members, Noam Chomsky, Jim Higginbotham and Donca Steriade have also made a substantial contribution. Noam pointed out a number of flaws in my thinking which had not occurred to me. Jim's help with the second and third chapters, especially in regard to the semantics was invaluable. Donca, undoubtedly the finnickiest member of my committee put more red marks on the first draft of this thesis than the rest of the committee combined: her careful attention to detail and her constant demands for clarification of my ideas often made me realize how much I had been taking for granted and how little I sometimes understood.

A number of other people have given me ideas and encouragement in the preparation of this thesis. Morris Halle read the chapter on Lexical Phonology and we spent a couple of pleasant afternoons discussing, often with great energy on Morris' part, the various implications of what I wanted to say. Tom Roeper gave me interesting comments on the second chapter. Alec Marantz read and commented on the first chapter; Alec was, in fact, one of the two people who first made me realize what it was I wanted to do for my thesis, since he (rightly I think) suggested that the problem of Bracketing Paradoxes was at the heart of an understanding of morphology. The other of the two mentors is David Pesetsky; although I lambast his approach to Bracketing Paradoxes in the first chapter, it is to

David that I owe the realization that there might, after all, be nothing in the idea of a separate word-formation component.

Turning now to the more general issue of graduate education, a number of people, some of whom have already been mentioned, gave a great deal to my educational experience at MIT. It was Noam who, in his lectures, and in my meetings with him, made me realize what it was about science that was fascinating; my understanding of science in general and linguistics in particular, meager though it may be, would be substantially more meager had it not been for Noam's insight. Morris Halle never let anyone get away with anything. his second semester phonology course was for me one of the more painful experiences of my tenure here. However, when my work did finally win his approval I knew for certain that I was learning to be a linguist. Ken Hale has been a constant source of encouragement. Since my work in the past has dealt with somewhat exotic languages he has always taken great interest in what I have been doing. To Jim Higginbotham I owe much of what I know about formal systems; his course in formal language theory, which we were all obliged to take our first semester here, has proven invaluable to me over the years. Bob Berwick of the Department of Electrical Engineering and Computer Science extended that knowledge into computational linguistics. Paul Kiparsky is probably the person most responsible for stimulating my interest in morphology. I also had some interesting discussions on topics that often extended beyond linguistics with Jim Harris and Wayne O'Neil. Finally, Jay Keyser, among all of the faculty deserves special recognition. While he was head of the Department--and undoubtedly now too that he has left for the higher planes of administration--his schedule was only equalled in hecticness by Noam's. Nevertheless, he was nearly always available for advice, which I certainly needed on an number of occasions. In many ways I doubt that I have met a wiser person than Jay.

My fellow graduate students have been indispensible for keeping the atmosphere here at MIT alive. To my classmates, Diana Archangeli, Mark Baker, Isabelle Haïk, Kyle Johnson, Juliette Levin, Margaret Magnus and Diane Massam, I am indebted not only for contributions to my intellectual life, but also to providing a social atmosphere which made the whole business of being a graduate student a heck of a lot more tolerable. Certain periods of one's life are often more difficult than others and for me the most difficult was probably the transition from my undergraduate life to my life here at MIT. My classmates were a great help in easing that burden. In previous and following years (and in one case in AI rather than in Linguistics), the following people in particular come to mind: Steve Abney, Andy Barss, Maggie Browning, Hyon Sook Choe, Jennifer Cole, Nigel Fabb, David Feldman, Margaret Fleck, Naoki Fukui, Alicja Gorecka, Tova Rapoport, Malka Rappaport, Marc Ryser, Doug Saddy, Mamoru Saito, Ur Shlonsky, Michele Sigler, Peggy Speas, Carol Tenny, Lisa Travis, Betsy Walli.

The following visitors and non-faculty/non-student types come to mind: Yosef Grodzinsky, Mary Laughren, Beth Levin, Jim McCloskey, Pieter Muysken, John Truscott, Tim Stowell. John Truscott seemed willing to give me a chess tournament at almost any time I felt like it, which was quite often while I was trying to avoid working on my thesis.

Two groups outside MIT deserve mention. First of all, my undergraduate teachers Sandy Chung, Ron Langacker, Margaret Langdon and David Perlmutter are the people I feel were most responsible for making it possible for me to come to MIT in the first place. Secondly, the Bell Labs crew Ken Church, Osamu Fujimura, Mark Liberman, Mitch Marcus and Janet Pierrehumbert gave me a valuable learning experience while I was an MTS at Bell Labs during the summer of 1984.

It would be remiss of me indeed were I not to mention the staff of the Department of Linguistics and philosophy, in particular Maggie Carracino and Nancy Peters, for their help with what must have often seemed like unreasonable demands at various times.

I am also grateful to the National Science Foundation for supporting me for three years as a Graduate Fellow.

Hoping, again, that I have not omitted anyone, I pass from my linguistics acknowledgees to my non-linguistics acknowledgees. My graduate life at MIT was made a great deal more enjoyable by the people who have at one time or another inhabited Ashdown House. The following deserve mention: Eugene Gath, Bob Gounley, Robert Granville, Bob Holt, Monty McGovern, Brian Oki, David Shera, and Mark Smith.

Finally, I wish to thank my parents and family for their encouragement over the years; it has not always been easy for them just as it has not always been easy for me, but we seemed to have pulled it off. The last person on my list is the person who made a great many things seem a great deal more worthwhile, Robin Vaughan. This thesis is dedicated to her.

Table of Contents

Abstract 2

Acknowledgments 4

Prologue 12

Chapter 1: On Bracketing Paradoxes: The Mapping Between S-Structure and PF 15

1.1 Bracketing Paradoxes 15
1.1.1 An Introduction 15
1.1.2 Siegel's Theory of Ordering 17
1.1.3 Allen's Treatment of un- 23
1.1.4 Fabb's analysis and an Argument against Fabb and Allen 27
1.1.5 Pesetsky (1979) 33
1.1.6 Williams (1981) 38
1.1.7 Strauss (1982b) 41
1.1.8 Guerssel (1985) 43
1.1.9 Kiparsky (1983) 45
1.1.10 Pesetsky again 51
1.1.11 A Summary so Far 65

1.2 The Mapping from Syntax to PF 66
1.2.1 Background: The model of grammar 66
1.2.2 The Condition on Mapping Between S-structure and PF 78
1.2.3 Simple cases of Bracketing Paradoxes 85

1.3 Handling Bracketing Paradoxes 88
1.3.1 Handling Pesetsky's data 88
1.3.1.1 No reordering in analysis 88
1.3.1.2 Idiosyncratically Interpreted Words 96
1.3.2 Kiparsky's Examples 97
1.3.3 The Affixation of -ity 103

1.4 More on the Mapping Principle 108
1.4.1 A curious piece of English morphology 109

1.4.2 Rebracketing of Separate Words 113
1.4.2.1 Consonant Mutation in Welsh 114
1.4.2.2 Other "Locality" Phenomena 120

1.5 On some cases which might seem to support morphological QR 123
1.5.1 Morphological QR in Navajo? 123
1.5.2 Other cases like Navajo 133

1.6 Conclusions and a Prospectus 137

Appendix: Some Further Consequences of the Mapping Principle 139

1 A Further Consequence of the Mathematical Properties of * and ^ 139
2 On Examples of Stratum I Prefixes outside Stratum II suffixes in English 141
3 Another Possible Example of Morpheme Doubling 143

Chapter 2: Syntax and Semantics of Productive Deverbal Morphology 147

2.1 Introduction 147

2.2 Background: Higginbotham (1985b) and Chomsky (1984) 149
2.2.1 Higginbotham's Semantics 149
2.2.2 Chomsky (1984) 159

2.3 The Formation of Agent Nominals 164
2.3.1 Introduction 164
2.3.2 -Er affixation 166
2.3.3 Projection, Projection and Theta Indexing 182
2.3.4 On Binding in -Er NPs 191
2.3.5 On Synthetic Compounding 195
2.3.5.1 The X-bar Status of Synthetic Compounds 195
2.3.5.2 Theta Marking within Minimal Projections 202
2.3.5.3 On the Application of the Projection Principle within Synthetic Compounds 207
2.3.5.4 Case Assignment in Synthetic Compounds 209
2.3.6 A short history of the study of Synthetic compounding and a comparison with Lieber (1983) 214
2.3.7 A short note on Chichewa synthetic compounds 225

2.4 The Syntax and Semantics of Derived Nominals 235
2.4 1 The Secret of NOM: the Morphological relationship between Verbs and Derived nominals 238
2.4.1.1 The Derived Nominal's Reference to an Event: Some Consequences 242
2.4.1.2 The Discharge of the Internal Theta Role 246
2.4.1.3 On the Discharge of the External Theta Role 257
2.4.1.4 Some Problems with and Predictions of the Model 264
2.4.1.5 On Synthetic Compounds with NOM 268
2.4.2 Not All Derived Nominals are Created Equal 280
2.4.3 A Discussion of Lebeaux (1984) 286

2.5 Concerning Self-analysis 289
2.5.1 -Er forms with self 293
2.5.2 Self- in Derived Nominals 297
2.5.3 Self- Prefixation to Nouns and Adjectives 301

2.6 Extensions of the Theory 306
2.5.1 A Short Discussion of Adjectival Passives 306
2.5.2 On Deadjectival Nominals 313

2.7 Conclusion: Overview and Prospectus 317

Chapter 3: Anaphoric Islandhood 326

3 1 Introduction 326

3.2 Simpson (1983) 329

3 3 On the Non-referentiality of Subparts of Words 335
3.3.1 Pronouns and NP anaphors 335
3.3.2 Non-Anaphoric-Islandhood with Submaximal Projections: Some examples 339
3.3.3 When Maximal Projections Do Occur within "Lexical" Constructions 345

3.4 A Summing Up 350

Chapter 4: Concerning Lexical Phonology 352

4.1 Introduction 352

4.1.1 A Brief overview of LPM 353
4.1.2 LPM as a Theory of Phonological Well-formedness 358

4.2 Against LPM as a Theory of Word-Formation 365
4.2.1 Morphological Blocking and Stratum Ordering 366
4.2.2 On the Stratum Ordering of Compounds 371
4.2.2.1 Inflection of Verbal Compounds 371
4.2.2.2 Inflection of Nominal Compounds 378
4.2.2.3 Malayalam Subcompounding and Cocompounding 382
4.2.2.4 Inflections within English Compounds 412

4.3 On the Phonology of Lexical Phonology 427
4.3.1 What are the Principles of Lexical Phonology? 427
4.3.1.1 Bracketing Erasure 428
4.3.1.2 On Cyclicity in Lexical Phonology 444
4.3.1.3 Structure Preservation 448
4.3.2 A Sketch of an Alternative to Lexical Phonology 459
4.3.2.1 Advantages of Stem/Word Morphophonology 460
4.3.2.2 Stem/Word Morphology and the Mapping Principle 468

Chapter 5: Some Final Points 472

5.1 Some Theoretical Issues 473
5.1.1 Morphological Blocking 473
5.1.2 What is a Component? 480
5.1.3 On the Projection Principle 487

5.2 Psychological Issues 491
5.2.1 On Productivity, Listing and the Separation of Word-Formation from the Rest of Linguistic Knowledge: What is the Relationship between the Theory of Morphology and My Knowledge of Words 491
5.2.2 On the Psychological Reality of the Lexical Level 499
5.2.2.1 Speech Errors 500
5.2.2.2 Language Games 501
5.2.2.3 Speech Recognition 502

5.3 Some General Conclusions and a Prospectus 505

Bibliography 509

"Words, words, words."

Hamlet, Act 2, Scene 2.

Prologue

Despite the fairly substantial body of research on the nature of the Word-Formation Component, there remains, to my mind, an assumption that has not at all been adequately justified, namely that such a component of the grammar exists. This thesis is an argument that the Lexicon does not exist, if by Lexicon we mean Word-Formation component.

There are a number of pieces of evidence which support this view, and I deal with these in turn. The first, which has also been discussed by Pesetsky (1985), is that if we are to seriously deal with the large number of cases which have come in the literature to be known as Morphological Bracketing Paradoxes, we must allow that word structure minimally consists of two levels of representation. The particular levels I argue for in the first chapter are syntactic and phonological levels of representation. I argue that a simple mapping relation holds between the two levels, which basically states that two entities **A** and **B** which are sisters in syntactic representation must have their phonological counterparts adjacent at phonological representation; see also Marantz (1984a).

The second point is that the syntactic representation of words does not need to be derived in a special Word-Formation Component, but rather can be

reduced to well-formedness conditions on syntactic representation, which conditions are simply familiar syntactic principles such as X-bar Theory, Binding Theory, Case Theory and Theta Theory; I show this in Chapter 2. Anaphoric Islandhood, too, which has been assumed to be a consequence of a separate Word-Formation Component is shown, in the third chapter, to be a trivial consequence of the fact that words do not (canonically) contain maximal projections, whereas maximal projections--i.e., arguments or predicates--must be present for anaphoric relations, such as pronominal coreference and binding, to take place.

Turning to lexical phonology, and in particular the theory of Lexical Phonology and Morphology, I show that lexical phonology does not really differ in its essential properties from what has come to be termed post-lexical phonology. In particular. Bracketing Erasure can be dispensed with entirely, Structure Preservation, depending upon how this notion is interpreted, is not necessarily true of lexical phonology, and cyclicity is neither characteristic of nor limited to lexical phonology. As far as phonological strata are concerned, these would seem to be reducible to specifications on the (phonological entries of) affixes; I suggest an alternative approach to Lexical Phonology which encodes phonological strata in the arguably more useful notions of phonological word and phonological stem.

Finally, in the last chapter I discuss some theoretical points including

the notion of "component of the grammar" and the question of morphological blocking. I also look at the psychological evidence, based on speech errors, secret languages and speech recognition, for the Lexical Level, which is the output of the lexicon in Lexical Phonology; I suggest that while this evidence may be correct--and is therefore evidence for any theory which allows for a level of representation, such as the Lexical Level or the Word Level, corresponding roughly to the Taxonomic Phonemic level--it is nevertheless clear that language processing does have access to deeper lexical levels of representation than just this level. The "lexicon" is therefore not a domain opaque to language processing as is implied by Mohanan's (1982) model of Lexical Phonology.

A basic theme running throughout this dissertation is that while one could well assume that a Word-Formation Component exists, it does not seem to be a necessary assumption and that, in fact, we can explain much more in many cases without this assumption. At best, then, the Word-Formation component becomes excess baggage, at worst a hindrance to explanation.

Chapter 1

On Bracketing Paradoxes: The Mapping Between S-structure and PF

1.1 Bracketing Paradoxes

1.1.1 An Introduction

It is an interesting characteristic of important problems in any field that they are very often the original problems which were posed at the field's inception, and yet which have eluded solution for generations despite repeated attempts. It is for this reason, I suppose, that (Western) philosophy is often termed "Footnotes to Plato": great Platonic questions such as the origin and nature of knowledge, and the referent of such ethical terms as "Justice" remain important topics of debate among philosophers to date.

There are substantially more modest problems which share with the great ones this characteristic of cotemporaneity with the founding of the field under which they are subsumed. One such modest problem is that of

Bracketing Paradoxes The morphological forms that would come to be known as Bracketing Paradoxes were noticed--and such forms were analyzed--in Siegel's (1974) dissertation, which was one of the first treatments of morphology within Generative Grammar. There have been various ways of dealing with the problems caused by these forms, either by making different assumptions about the nature of particular affixes (Siegel, 1974; Allen, 1978; Selkirk, 1982; Fabb, 1984), by arguing for the non-application in certain cases of otherwise supposedly general morpho/phonological principles (Kiparsky, 1983), by arguing for special notions of word-structure interpretation (Williams, 1981a, 1984; Strauss, 1982) or, finally, by arguing that there is more than one level of representation for word structure, as indeed for other linguistic entities (Pesetsky, 1979, 1985; Marantz, 1984a,b,c; Sproat, 1984). This latter approach is the one which will be taken in this dissertation.

Insofar as I will provide a solution to this problem, it will turn out that Bracketing Paradoxes are in no sense paradoxical and hence that the problem always has been a trivial one. In another sense, however, the problem has always been significant since providing a solution could in principle--and I shall argue in fact does--give a very deep insight into the nature of grammatical organization. In particular, based partly on the evidence from Bracketing Paradoxes, I shall end up arguing that there is no lexicon, if by "lexicon" we mean the component of the grammar which is responsible for word formation. I shall, of course defend this particular

conjecture in the remainder of the thesis.

In this chapter, then, I shall provide a characterization of the problem of Bracketing Paradoxes, which I feel gives a natural and satisfactory solution, and at the same time forces a particular view of the organization of the grammar. I start by tracing the history of the problem.

1.1.2 Siegel's Theory of Ordering.

Siegel, in her (1974) dissertation, put forward the hypothesis that derivational word-formation in English consists of two levels of affixation and that these two levels are ordered with respect to one another. These two levels correspond to Chomsky and Halle's (1968) two types of affixes, the '+' affixes and the '#' affixes. '+' affixes are usually what has been termed latinate affixes--affixes such as _-ity_, _in-_ and so on. '#' affixes, on the other hand, are native affixes such as _-less_ and _-ful_. Following SPE, she notes that the affixation of these two classes of affixes results in different phonological effects. For example, the latinate affixes, which she calls Class I, generally result in stress shift, whereas the non-latinate '#' boundary affixes, generally do not result in stress shift. Examples are given below:

(1)

grammatical	grammatical+ity	grammatical#ness
special	special+ity	special#ness
equal	equal+ity	equal#ness

In the above words, the primary stress is marked by underlining the appropriate vowel. The (by now very familiar) contrast between the '+' affix -ity and the '#' affix -ness is that the former shifts stress in the word to which it attaches whereas the '#' affix -ness does not.

Siegel also noted that '+' affixes such as -ity attach to stems as well as to fully formed words, whereas '#' affixes attach only to words. For example, -ity attaches to stems such as prosper in prosperity, whereas -ness can do no such thing: *prosperness. Her central thesis was that the two fundamental properties of '+' affixes, namely stress shift and affixation to non-words were related, and were in particular a consequence of the ordering hypothesis stated as follows:

(2) Ordering Hypothesis (Siegel, 1974)

A. In English, Class I affixation precedes Class II affixation.
B. The cyclic stress assignment rules follow Class I affixation and precede Class II affixation.

The elaboration of this hypothesis, in the work of Allen (1978), Pesetsky (1979), Mohanan (1982) and Kiparsky (1983a, etc.) has led to what is now termed Stratum-Ordered Morphology, a topic to which I will return in some depth in the fourth chapter of this dissertation.

The Ordering Hypothesis obviously has the desired effects for the data in (1) since -ity, being a Class I affix, will attach to a word (or a stem) and will be in the domain of stress assignment, which postcedes its affixation. -Ness on the other hand is Class II and comes after stress

assignment and therefore obviously cannot affect stress.

As Siegel notes, the organization of the morphology implied by the Ordering Hypothesis makes a prediction besides the predictions it already makes concerning stress placement: Class I affixation ought to consistently appear inside Class II affixation, but the reverse should never be true. For the most part, this seems to be correct[1]. To take a simple example, there is a contrast between the two negative prefixes non- and in-. Non- is Class II whereas in- is Class I, and this correlates with the fact that one can find in- inside non- but not non- inside in-; non-illegible vs. *in-non-legible, though non-legible is fine. This prediction, of course, should carry over to cases where one of the affixes is a prefix and the other is a suffix; that is, we should never (or at least rarely), according to the Ordering Hypothesis find words of the form [[Prefix X] Suffix], where the bracketing is arguably as given, and yet where the Prefix is Class II and the suffix is Class I.

---

1. Though see Aronoff (1976) and Aronoff and Sridhar (1983) for arguments that it is not. Again, I shall be looking more closely at the correctness of Stratum-Ordered Morphology in Chapter 4, and I shall discuss Aronoff and Sridhar's ideas on this there. I should point out at this juncture, though, that the ultimate correctness of Stratum Ordered Morphology has no effect on the discussion to be given below concerning the paradoxical status of forms like ungrammaticality. Irrespective of one's theory of morphology, there are reasons for saying that -ity is bracketed inside un- as far as the phonology is concerned: Aronoff and Sridhar end up with this conclusion despite their arguments against Stratum-Ordered Morphology.

Yet there are examples of precisely that form, and even in Siegel's day they must have seemed somewhat paradoxical. A famous class of cases are the nouns of the form un-X-ity where X is an adjective and where the bracketing looks as if it must be [[un X] ity]; the reason for this assertion comes from the fact that un , as Siegel noted, attaches productively only to adjectives (we are not, of course, talking about the verbal prefix un- which is arguably different from the adjectival prefix.) We shall argue for the correctness of this claim about the subcategorization requirements of un- below.

Now, -ity is a Class I suffix, and the question of whether un-X-ity forms are paradoxical depends upon the status of un-; if it is Class I then there is no problem since the Ordering Hypothesis says nothing in general about the affixation of Class I to other forms derived by Class I affixation. If, on the other hand, it is Class II, then we have a problem since the un-X-ity forms look like counterexamples to the Ordering Hypothesis. Un- does indeed seem to have some characteristics of Class II affixation; first of all it is a native affix and does not appear to require that the adjective to which it attaches be latinate. So we have forms such as unhappy but not *inhappy, and unmade but not *inmade~immade. In- also attaches apparently to stems, as well as to words, whereas un- does not (though there are exceptions, c.f. *couth versus uncouth). The following examples are from Allen (1978):

(3)

*ert	inert	*unert
*placable	implacable	*implacable
*trepid	intrepid	*untrepid
*sipid	insipid	*unsipid
*maculate	immaculate	*unmaculate

Yet Siegel had what appeared to be good arguments to suggest that un- was Class I, and hence that adjectives of the form un-X-ity were unproblematic for the Ordering Hypothesis. For one thing, she notes that un-X-less forms are bad:

(4) *unharmless, *unspeechless, *unhopeless, *untasteless.

Now, -less is a Class II suffix--note that we have already shown that it does not shift stress. If un- were a Class II prefix then we would expect it to attach to these adjectives, and yet it does not. If it is Class I, however, we would not expect the forms in (4) to be good, as indeed they are not. So, by characterizing un- as Class I, we simultaneously rule out forms such as *unharmless and predict the occurence of other forms such as ungrammaticality.

Furthermore, certain un-X forms undergo a rule of destressing on un- when the un- precedes a stressed syllable in the adjective: unable, unaided, unopen. According to Siegel, Destressing does not occur with Class II prefixes: one such example is the word disinclination which Siegel claims has the structure $\#[\text{dis}\#]_P[\#\text{inclination}\#]_N\#]_N$. In terms of the Ordering Hypothesis, Destressing, as part of the block of cyclic stress

assignment rules of English, applies after Class I affixation, but before Class II affixation. It will therefore necessarily be blocked with Class II affixes. Yet un- seems to be destressed in the forms listed above, and this suggests that un- is Class I.

Siegel would thus appear to have nipped in the bud a problem for the Ordering Hypothesis. It should be noted, before we go on to discuss Allen's arguments against the Class I-hood of un- that Siegel's analysis has in effect been reintroduced into the literature by Selkirk (1982). Selkirk claims that ungrammaticality and other such Paradoxical forms can be handled if we allow un- to behave as a Stratum I (i.e., Class I) prefix. Selkirk does not deny the evidence of Allen (1978; and also Kiparsky, 1983a; Mohanan, 1982), that un- has Stratum II (Class II) properties, and she suggests in fact that un- is both Class I and Class II. It will become clear that this approach is unnecessary; un- will be Class II, and still be able to attach "inside" Class I affixes. In any event, Selkirk's approach is less desireable than Siegel's: even though it will be shown momentarily that Siegel's characterization of un- as Class I is wrong, at least Siegel's theory had the property that it made strong predictions about the kinds of words we should expect to find. Characterizing un- as both Class I and Class II, is a weaker theory of morphology than we might like to have, unless we have very strong evidence that such a weakening is necessary.

1.1.3 Allen's Treatment of un-.

As noted above, there are problems for the analysis of un- as a Class I affix. There is no doubt that that would be a necessary conclusion if one wanted to account for the existence of many un-X-ity forms given that un- does not productively attach to nouns. However, a major contribution of Allen's work to this issue was the demonstration that this conclusion is nevertheless untenable.

We have already seen one of Allen's arguments in the last section, namely that in-, which is unquestionably Class I--it is latinate after all--attaches to stems whereas un- only attaches to fully-formed words. Another important blow to Siegel's analysis is the fact that although un- does not attach to adjectives of the form X-less it does attach to adjectives which are formed by other Level (Class) II processes (Allen, p. 31):

> But the argument does not hold. If it did, then adjectives of the following types should not exist.
>
> (31) $[un\ [[X]_V\ ing]_A]_A\ [un\ [[X]_N\ ed]_A]_A\ [un\ [[X]_N\ ly]_A]_A\ [un\ [[X]_N\ ful]_A]_A$
>
> But all these adjectives are well-attested types; unbending, unrewarding, unprecedented, unterraced, unmanly, unfriendly, uneventful, uncheerful. The examples in -ful are particularly striking because of their parallelism to the non-existent cases in -less. There are inumerable words of the type un-X-ful, for which we must assume a bracketing identical to that of the

non-existent un-X-less type; i.e., [un [X-ful]$_A$]$_A$, uneventful, uncheerful, unsuccessful, unfaithful, unfruitful, unmindful, unmerciful, unthoughtful, untruthful, unresentful, ungrateful, unhelpful, unforgetful. It is also quite clear that -ful is a Level II, word-boundary, suffix. Siegel actually presents a lengthy argument to this effect[2].

The upshot of this is that un- cannot be a Class I affix, since otherwise one cannot account for the extreme productivity of the affixation of un- to adjectives which must be derived by Class II affixation. The fact that un-X-less forms are bad is attributed by Allen to the following condition (=her 59, attributed by her to Zimmer):

(5)
Condition on un- prefixation (Generalized): Un's base may not have negative content

This accounts not only for the *un-X-less forms, but also for the non-existence of the following forms (all examples from Allen):

(6)
*undishonest, *undisloyal, *unmalnourished, *unhorrible.

Siegel's other argument for the Class I-ness of un-, namely the destressing of un- directly before stressed syllables is claimed by Allen to be based on a misanalysis of the data. The fact is that un-, unlike clear Class I affixes is never reduced to 0-stress by the destressing rule. In particular, Allen claims there is no difference between the

---

2. Though see Chapter 4 where I argue that the evidence for this property of -ful is not so clear.

stressing on un- in the A forms in the following example, where the destressing environment is met, and in the B forms where it is not (= Allen's 38):

A	B
unable	unexcitable
unopen	uneventful
unordered	unkind
unelevated	unofficial
uneven	unoriginal

If this is correct, of course, then there is no phonological evidence for the Class I status of un-

Of course, all of this leaves us in a quandary as to what to do about the un-X-ity forms, which now seem to be paradoxical indeed. As Allen notes, if we assume the Ordering Hypothesis, the Class I-ness of -ity, the Class II-ness of un- and the subcategorization restriction on un- that it can only attach to adjectives, then we seem to be forced into the conclusion that un-X-ity forms must on the one hand have the bracketing [un [X-ity]] (since Class II affixes must be outside Class I affixes), and on the other hand the bracketing [[un X] ity] (since un- attaches only to adjectives and X-ity forms are deadjectival nouns, not adjectives.)

The solution that Allen proposes is that one of the assumptions, namely that un- attaches only to adjectives, is incorrect. She notes first of all that un- does not productively attach to X-ity forms (p. 33):

> Even a cursory examination reveals that un- is completely unacceptable in the vast majority of cases with X-ity; e.g. *unsobriety, *unagility, *unpropriety, *unopacity, *unsincerity, *uncivility, *unpurity, *unfertility, *unregularity, *unhospitality.

And as she shows further un- apparently does attach to nouns: uninvolvement, unemployment, unfulfilment, unacceptance.

If this is a correct assessment of the status of un-X-ity forms, of course, then these forms do not constitute Bracketing Paradoxes; the bracketing required by the selectional requirements of the affix un- lines up exactly with that required by the Class status of the affixes. It will turn out, perhaps unsurprisingly, that even if Allen s assessment is correct, then we will not have buried the Bracketing Paradox issue since there are other examples, to be discussed below, which do not involve the affixation of un- and -ity, but which nevertheless have the same apparent structural paradoxes.

Nevertheless, I shall show that Allen's characterization is incorrect, and that un-X-ity forms are paradoxical for a theory of morphology which only allows for one structure for words, in particular for the Stratum-Ordered Morphology discussed in Allen, and furthered in the work of Kiparsky and Mohanan. Before I demonstrate this, however, I digress to discuss Fabb's analysis of un-X-ity forms, which is essentially the same as that of Allen.

1.1.4 Fabb's analysis and an Argument against Fabb and Allen.

Fabb (1984), who is concerned with the partial reduction of Stratum Ordering in Phonology to principles of selection on affixes--a topic to which I return in a later chapter--notes that un-X-ity forms are problematic for his theory in the following way. By assuming that latinate affixes such as -ity select for a feature such as [+latinate] on the stem to which they attach and that they themselves carry that feature, and by further assuming that native affixes such as un- are not restricted in their attachment, but do not carry the feature [+latinate] he is able to derive the fact that latinate affixes occur inside native affixes, but that native affixes do not occur inside latinate affixes. He argues convincingly that these selectional restrictions, and also the feature [+latinate] (as also noted by Aronoff, 1976), are necessary quite independently of the question of whether or not there are Strata of word formation as claimed by Allen, Kiparsky and Mohanan. As already noted, I will be returning to the question of the status of Stratum-Ordering in a later chapter and therefore will not dwell on this issue here. For the present purposes it is sufficient to note that given Fabb s analysis, un-X-ity forms are as problematic for him as they were in the model Allen was assuming. That is, if un- must attach to adjectives, then a form such as ungrammaticality must be bracketed as follows: [[un grammatical] ity].

But given Fabb's assumptions, it turns out that ungrammatical would end up not bearing the feature [+latinate] and therefore -ity would not be able to attach to it.

So Fabb ends up making the same assumption as Allen does, namely that un- attaches to nouns as well as to adjectives. He actually makes a more specific claim than that, stating that un- attaches to words which refer to states, whether they be adjectives or nouns. He gives the following examples of words where un- has attached unequivocally to a noun (= his 5.198):

(7)
un-dismay, un-concern, un-balance un-belief, un-embarrassment, un-involvement, un-ostentation, un-acceptance, un-alarm, un-promise, un-bias, un-charity, un-employment, un-fulfilment, uncompassion, un-being.

To this list could be added such forms as unhealth and unease which are acceptable for at least some speakers, including myself.

Nevertheless, despite the obviously large number of cases where un- clearly attaches to nouns, I will argue that neither the more specific thesis presented by Fabb, nor indeed the more general version of the analysis presented by Allen is satisfactory. I will not attempt to deny that there are cases where un- attaches to nouns; to do so would be absurd in light of the examples given above. Rather, I shall deny that un- productively attaches to anything but adjectives.

There is, despite Allen's claims to the contrary, a great difference between un-X-ity forms and the other un-N forms which she gives. The fact of the matter is that while we cannot predict whether a given noun will take un- as a prefix, we can say that, in a large class of cases (see section 1.3.3 for a more detailed treatment), if un-X occurs, where X is an adjective, and if X-ity also occurs, then un-X-ity will occur. Now, it is a fact that not all adjectives accept un-. Some examples are given below:

(8)
??unagile, ??unopaque, *unsincere, *unpure. *unfertile, *unregular

Some of these forms (pure, fertile, regular, sincere). have negated forms with the latinate prefix in-; the non-existence of un-A forms for these adjectives may thus be explained by morphological blocking (see in particular, Aronoff, 1976). The other two (??unagile and ??opaque) do not seem explainable in this way, but in any event, they seem bad. It is not my purpose at this point to investigate the class of adjectives to which un- attaches; for a fairly nice description, see Aronoff (1976, p. 63; and see below, section 1.3.1.1).

The astute reader will have noticed that the adjectives listed as bad here are the base forms for many of the -ity words which Allen lists as not taking un-. Thus, to repeat the examples:

(9)
*unagility, *unopacity, *unsincerity, *unpurity, *unfertility, *unregularity.

A reasonable theory of morphology should relate the non-existence of the forms in (9) with the non-existence of the corresponding adjectives. But this is what Allen's, and subsequently Fabb's, theory precisely fail to do. In fact, far from being problematic for the claim that un-X-ity forms are bracketed [[un X] ity] as Allen argues, the lack of forms such as those in (9) is precisely what would be expected under such a theory, given that the un-A forms are lacking. On the other hand, these forms are problematic for Allen's and Fabb's theories precisely because there is no reason, given the fact that un- attaches both to adjectives and nouns, why un- should fail to attach to any of the base -ity forms in (9)[3]. The forms are especially problematic for Fabb since all of the nouns refer to states and therefore ought to be affixable by un-. Yet they are not.

Several of the other examples given by Allen are irrelevant as far as a theory where un- only attaches productively to adjectives is concerned, though they remain problematic for Allen. These are *unpropriety, and *unhospitality. The question here is, what is the base adjective in these

---

3. Note that the absence of many of these forms can be explained by morphological blocking by the corresponding in-----ity forms: so, insincerity will block *unsincerity. But this is missing the point insofar as given that un- attachment to Ns is independent of its attachment to adjectives, there is no reason why we could not find pairs of the form in-A, un-A-ity; i.e., why could un- not attach to sincerity in preference to in- independently of the fact that the related adjective is insincere. Allen and Fabb's analysis ought, if anything, to predict complete independence for un- or in- of an adjective and its -ity form, which is not characteristically found.

forms to which -ity is attaching? It is certainly not *propri, or *hospital since there are no such adjectives. (If the base adjective of propriety is proper, then this example falls under the cases discussed above since, again, *unproper is bad, being blocked by improper.) It should be obvious that a theory where un- attaches productively only to adjectives makes no claim that it will attach productively to all nouns ending in -ity. In fact it makes precisely the opposite claim. The forms are problematic for Allen and Fabb, however, since again, they are both nouns (which refer to states) and there should be not reason off-hand why un- should fail to attach to them.

Where un- does attach to adjectives however, it is often the case that there is a corresponding -ity word with un-:

(10)

uncomputable	uncomputability
unparsable	unparsability
unrecognizable	unrecognizability
unthinkable	unthinkability
unpronounceable	unpronounceability
ungrammatical	ungrammaticality
unhomogeneous	unhomogeneity
unrhythmical	unrhythmicality
unpractical	unpracticality
unveracious	unveracity

These are my own judgments, in many cases. This is not to say that such is always the case and a more complete treatment is given in section 1.3.3. Some of the *un-X-ity words listed by Allen constitute such exceptions: *unsobriety despite unsober, and *uncivility despite uncivil. In many such

cases, however, we may again be dealing with blocking; c.f., the existence of incivility and insobriety. Often, however, especially in X-able adjectives, the generalization is quite powerful. This is a significant fact because it is precisely what is expected given the restriction on un- attachment to adjectives, and the assumption that in un-X-ity adjectives, the un- is somehow, attached inside the -ity. None of this is expected on Fabb's or Allen's account, however, since the attachment of un- to a noun is quite independent of whether or not it attaches to the corresponding adjective.

What, then, do we say about the cases where un- clearly does attach to a noun? I claim that these cases constitute an irregular set and are not instances of the regular affixation of un-. This should not seem like a surprising conclusion: many affixes regularly subcategorize for words of a particular class, but attach, in the marked case to words of a different class. To take just one example, agentive -er attaches only to verbs, but there are quite a few examples where it attaches to nouns with an agentive meaning:

(11)

footballer	*to football
?baseballer	*to baseball
carpet-bagger	*to carpet-bag (except, perhaps via back-formation)

Fabb may indeed be correct in conjecturing that the property of referring to a state is relevant even in the irregular affixation of un-: and this

may, in turn, be parasitic off of the fact that un- productively attaches to adjectives, which usually refer to states.

It seems, however, that un- affixes productively only to adjectives, and that we want to derive the existence/non-existence of many un-X-ity forms from the existence/non-existence of the corresponding un-X forms. Needless to say, this ends us up with precisely the same paradox as the one which we started out with since we still want to maintain the Class II-ness of un- for the reasons outlined in Allen (1978). But, as noted above, and as we shall see in the next section, the existence of morphological Bracketing Paradoxes is clear from examples which have nothing to do with this rather particular problem in English morphology.

1.1.5 Pesetsky (1979)

The treatment of Russian phonology which Pesetsky provides in this paper is one of the seminal analyses leading to the theory of Lexical Phonology as discussed in Mohanan (1982) and Kiparsky (1983a). In this paper he provides an analysis in which affixation is concomitant with phonology in that newly affixed forms are submitted to the phonology as soon as they are created. In addition to the importance for the theory of phonology of the analysis of Russian presented in this paper, Pesetsky also provides important data which are paradoxical in the same way as we have seen un-X-ity forms to be, thus demonstrating that Bracketing Paradoxes are real

and not merely an artifice of one problem of affixation in English.

The first examples Pesetsky notes are from Russian. In Russian as in many languages, there are many verbs which consist of a compound of a verb root with one of a set of preverbal particles. Now, insofar as the Prefix-V structure does not have a semantic interpretation compositional of the meaning of the prefix and the verb (c.f., the analogous cases in English, such as forget, forgo, withdraw, understand), the prefix and the verb do form a lexical item. Therefore, when inflectional suffixes are added one would expect that they would be bracketed outside the prefix and the verb, for the purposes of semantic interpretation, as follows:

(12)
[[Prefix V] inflectional-affixes]

But there is phonological evidence that the bracketing must go the other way, namely

(13)
[Prefix [V inflectional-affixes]]

To take a specific example, the prefix podŭ- 'up to' and the verb stem žĭg 'burn' combine to form a verb meaning 'set fire to . Now this meaning is, to be sure, related to the meaning of the base verb, but it is not predictably related. As such, we would expect that, for the purposes of semantic interpretation, the prefix and the verb would be bracketed together as a constituent and interpreted as a single unit idiomatically. This would yield the bracketing [[podŭ žĭg] inflectional-affixes]. However

there is evidence that the bracketing for the intents and purposes of the phonology must be [podŭ [žĭg inflectional-affixes]]; the argument here has to do with the correct statement of the rule of yer-lowering, which Pesetsky argues can be stated in a maximally simple way if we assume that prefixes are indeed external to the verbal root plus its inflectional affixes. We thus appear to have a bona fide paradox.

Pesetsky argues, however, that morphological bracketing paradoxes are by no means uncommon. He cites an example from Warlpiri which involves, again, verbal prefixes and inflectional suffixes, and is therefore completely analogous to the Russian case. So, the word pardi-mi means 'arise+NONPAST' and tirl-pardi-mi means 'open (as of an eye)+NONPAST. But even though this obviously suggests the bracketing [[tirl pardi] mi], Nash (1979) provides phonological evidence from cyclic stress assignment that the bracketing must be [tirl [pardi mi]].

Closer to home, there is the problem of un-Comparatives in English. The comparative affix -er is restricted phonologically in the types of adjectives to which it can attach. The generalization is that it attaches to maximally bisyllabic adjectives (examples from Pesetsky, 1985):

(14)
a. blacker, softer, riper, tougher, truer....
b. happier, heavier, luckier, kindlier, pleasanter....
c. *eloquenter, *irrascibler, *importanter.....

Actually, there are bisyllabic adjectives to which -er will not attach, for

example:

(15)
*directer, *activer, *complexer....

Pesetsky suggests that the restriction is probably stronger, and that _-er_ will not attach to bisyllabic adjectives with a heavier second syllable (though note that this does not apply to _active_). In any event, as he notes, whatever the precise characterization of the phonological restriction is, it is clear that _-er_ will not attach to trisyllabic adjectives. Given this, it is obvious enough what needs to be said about the following forms:

(16)
unhappier, unpleasanter, unluckier...

Since the addition of _un-_ turns the base bisyllabic adjective into a trisyllabic adjective, the bracketing of the words must be as follows:

(17)
[un [happi er]], [un [pleasant er]], [un [lucki er]]

But, again, there is good reason to believe that this cannot be the right structure as far as the semantic interpretation of the words is concerned. As Pesetsky notes, for example, _unhappier_ does not mean 'not more happy' but rather 'more not happy'. That is, if we say that John is not more happy than Bill, then it might be the case that John is equally happy with Bill; if John is unhappier than Bill, then it must be the case that John is happy to a lesser extent than Bill. This, of course, implies

that the bracketing of the words in (16) must be:

(18)
[[un happ'] er], [[un pleasant] er], [[un lucki] er]

Again, we have a paradox.

So there are copious data which suggest that Bracketing Paradoxes are real, and not limited to a small class of morphological processes in English. The problem is, rather unsurprisingly, what to do about them.

There have basically been two methods for handling the Paradoxes of the kind discussed here, and for the sake of discussion we can dub them "One Structure" analyses and the "Two Structure" analyses. One-structure analyses have been discussed by Strauss (1982), Williams (1981) and most recently by Guerssel (1985), whose proposal is by far the most interesting of this class. Two-structure analyses have been proposed by Kiparsky (1983c), Pesetsky (1985), and by Marantz (1984c). The analysis to be presented here will also be a two-structure analysis.

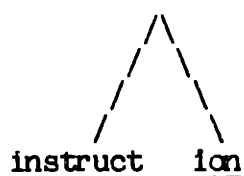
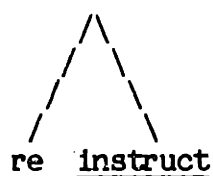
Pesetsky, in his original paper, provides what would be termed under this classification a One-structure analysis. He assumed that the structure dictated by the phonological analysis was the correct structure and that complex verbs were built up by the Level-ordered morphology accordingly. Rules of semantic interpretation of the following form would then be invoked (= Pesetsky's (126)):

(19)
Abstract Meaning of Prefix-Abstract Meaning/Root--->Idiosyncratic Meaning

In the next few sections I briefly discuss the other One-structure analyses. I then go on to discuss the proposals of Kiparsky and Pesetsky, which are more similar to my own.

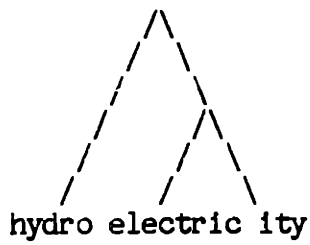
1.1.6 Williams (1981)

In this paper, Williams gives a characterization of the notion of lexical relatedness in terms of his notion of head. He defines head for the purposes of (English) morphology as follows (p. 248):

> In morphology, we define the head of a morphologically complex word to be the righthand member of that word. Thus, the head is italicized in (9a,b):
>
> (9)
> a. 
>
> b. 
>
>
> Call this definition the Righthand Head Rule (RHR).

Williams examines cases of the following form, which he terms "Relatedness Paradoxes":

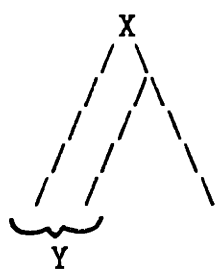
(20)



The point here is that hydroelectricity, although it has the structure indicated in (20), nevertheless is related to hydroelectric. Needless to say, these are precisely the Bracketing Paradoxes which we have been discussing up till now.

As, Williams notes, the structure of all of these paradoxes is as in (21) below:

(21)



That is, X is "related" to the word that would consist in the concatenation of the two morphemes bracketed here as Y.

How do we get relatedness to fall out if we assume that the structure of X is not somehow reanalyzed? Williams makes the following proposal, referring to the structure given in (21) above:

What we need is a definition of "related" which will let X and Y be related in such a structure. Suppose we said that X could be related to Y if Y would be the result of removing the head of X--then X and Y could be related to Y in such a structure. For example, the head (or, one of the heads) of Godel numbering [sic] is -ing--if we remove this head, we get Godel number.....

We will want to generalize this slightly. To relate macroeconomist to macroeconomic, we will want to say that two words can be related if one can be derived from the other by varying one of the heads:

(47)



We will also want to relate macroeconomic to microeconomic, where the nonheads are different. Suppose we define the nonhead of a word as follows:

(48) Nonhead: the highest left branch of a word.

...We may then defin relatedness as follows:

(50) X can be related to Y if X and Y differ only in a head position, or in the nonhead position

I have nothing in particular to say about Williams' proposal; it works well enough in covering the data he has set out to cover[4]. It does involve

---

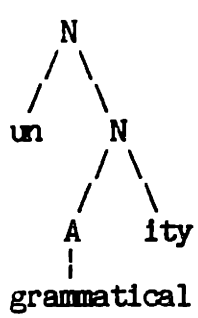
4. Though, as Strauss (1982a) notes, there is a problem with the definition of subcategorization if we assume that ungrammaticality has the structure:

some definitions ("nonhead", "related") which I would claim are unnecessary, or at least ought to follow from more general principles. Also, Lieber (1980) has argued against Williams' notion of "head" of a word, and the reader is referred to her discussion on that. It will turn out, in any event, that the proposal that I will present will cover a broader range of data than Williams' can including some facts which heretofore have not been considered to be part of the same phenomenon as Bracketing Paradoxes. Insofar as this will be a correct amalgamation, my proposal will have legitimately broader coverage, and hence be an improvement over Williams'.

1.1.7 Strauss (1982b)

I will have even less to say about Strauss's proposal than I had to say about Williams'. Strauss (Chapter 2) notes that forms such as

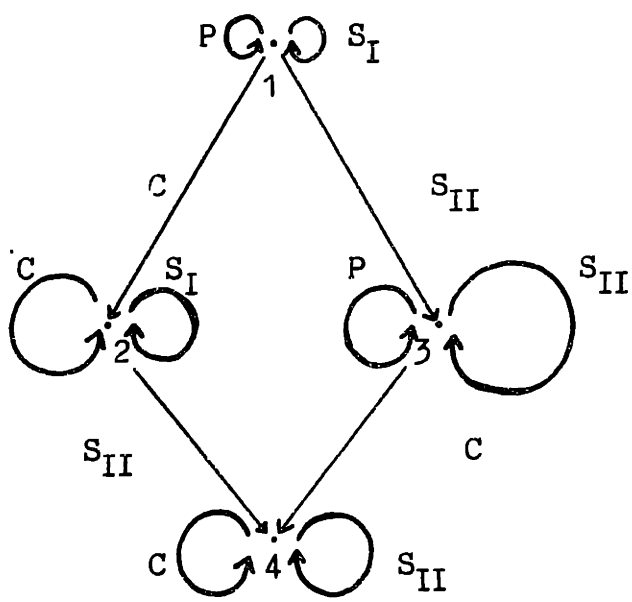
---



We not only have to account for why ungrammaticality is related to ungrammatical, but also why un- is allowed syntactically to attach to the noun grammaticality despite its selection for adjectives.

ungrammaticality are problematic for a stratum-ordered morphology. In part to account for this, he proposes a model of morphology which is stratum-ordered, but which has prefixation and suffixation as separate processes. This allows for prefixes to be adjoined to stems independently of the adjunction of suffixes. In particular, the model allows for the prefixation of "compounding elements"--namely true words and Class II prefixes, which Strauss argues have a "compound status"--prior to the affixation of Class I suffixes. His finite state diagram for the model is as follows (where P=prefixation, C=compounding, S_I=Class I suffixation, and S_{II}=Class II suffixation):

(22)



Note in particular that placing a loop for S_I at State 2 allows for the

suffixation of Class I suffixes after the prefixation of compounding elements, including Class II prefixes.

Nonetheless, while this model undoubtedly derives the right results, it appears to give little insight into the nature of the problem; we could as easily imagine a state diagram with a totally different set of specifications. Besides, it should not be necessary to stipulate that prefixes and suffixes behave as if they are independently affixed. Indeed, this independence of behavior will follow from the account of Bracketing Paradoxes to be presented here.

1.1.8 Guerssel (1985)

Guerssel, following work of Brame (see in particular, Brame 1984; and also Guerssel, 1983), argues for a theory of morphology in which suffixation is conceived of as applying a function to a base, and that successive suffixation is simply the composition of such functions. He notes that in English prefixes enjoy a different status from suffixes, in that while the latter have head-like properties, the former do not (see, again, Williams, 1981a). From this he argues that whereas suffixes select for their bases, prefixes are selected for <u>by</u> the heads of their bases. In other words, suffixes are functions which take arguments which are their bases, whereas prefixes are the arguments which are selected for by the head of the word.

To see how this all works, consider a bracketing paradox such as underivability. The components are un-, which is an argument which Guerssel notates as A^{un}; derive is a verb, which, being a syntactic argument taker itself, is a function, notated F^{V}. Finally, -able and -ity are functions notated F^{able} and F^{ity} respectively. Now, given that adjectives in -able apparently select for un---note the fact that un- productively attaches to any such adjective--and also that -ity selects for F^{able}, we can derive the fact that underivability is well formed. First, derivable is well formed since -able selects for verbs:

(23) $F^{able} \circ F^{V}$

Next, underivable is well formed insofar as compositions of the form $F^{able}\ F^{V}$ select for A^{un}:

(24) $(F^{able} \circ F^{V})(A^{un})$

Finally, underivability is well formed since -ity selects for -able:

(25) $(F^{ity} \circ F^{able} \circ F^{V})(A^{un})$

The satisfaction of the requirements of ity follows from the associativity of composition--which is stipulated as a part of Brame and Guerssel's theory--since F^{ity} and F^{able} may be composed together before composition with the other functions:

(26) $(F^{ity} \circ F^{able}) \circ F^{V}(A^{un}) = F^{ity} \circ (F^{able} \circ F^{V})(A^{un})$

The associativity of composition in Brame and Guerssel's theory is mirrored, as we shall see, in the associativity of concatenation (^) in the theory I shall be presenting below. Guerssel's approach is therefore quite similar in spirit to mine. On the other hand, the theory of grammar which Guerssel is assuming (namely that of Brame) is so utterly different from the one which I am going to be assuming that serious comparison of the two is difficult.

1.1.9 Kiparsky (1983)

We now turn to a discussion of proposals which are similar to the one I intend to make. What Kiparsky suggests is that there is a process or rule of rebracketing which takes place in the lexicon and which rebrackets structures which are ill-formed into structures which are well-formed (assuming such a reanalysis exists). The principles which govern this rule are as follows (pp. 24-5)

(27) 1. A form of the Projection Principle holds in the morphology. It stipulates that <u>subcategorization requirements of affixes must be met at every level.</u>

2. Reanalysis of morphological bracketing is freely permitted at any point in the derivation, subject to the above "Projection Principle."

He then goes on to show how ungrammaticality would be formed. First of all at Stratum I, affixation of -ity would take place yielding a structure of the following form:

$[\text{grammatical}_A \text{ ity}_N]$

At Stratum II, affixation of un- takes place As already noted un- attaches to adjectives and not to nouns, so the bracketing obtained by attaching un- to grammaticality cannot stand as is. Rather it must be reanalyzed so as to satisfy the morphological Projection Principle into the following structure:

$[[\text{un grammatical}_A] \text{ ity}_N]$

Now the subcategorization requirements of both un- and -ity are both satisfied and the structure is ruled in.

This approach, however, is not unproblematic in Kiparsky's framework. Although he is assuming that there are, in effect two structures for Bracketing Paradox cases such as ungrammaticality--i.e., the one generated by the blind application of morphological rules according to the Stratum Ordered morphology, and the one given by the rule of reanalysis, he is nevertheless assuming that there is, at any given time, only one structure for the word. That is, the rule of reanalysis looks at the same structure as the rules of the phonology look at and this is where the problem lies, because Kiparsky also assumes that there is a general convention which

applies in the lexicon erasing internal brackets at the end of every stratum of word-formation. This convention is termed the Bracketing Erasure Convention[5], and is given below in (28):

(28) Bracketing Erasure Convention
Internal brackets are erased at the end of every level. (= stratum)

According to this principle, after a word, such as grammaticality, is formed at Stratum I in the morphology the brackets which delimit the morphemes concatenated at that Stratum are erased. Thus, when further affixation is done at subsequent Strata there is no way for those affixation processes, or the rules concomitant with them, to distinguish a morphologically complex word (such as grammaticality) from one which is not morphologically complex (such as happy). How then can reanalysis occur in a form like ungrammaticality at the Stratum at which un- is attached, if there should be no principled way to see that grammaticality is morphologically complex (and hence contains an adjective to which un- can attach) at that Stratum? Kiparsky, of course, realizes that this is a problem, and makes the following proposal (p. 25; again, for "level" read "stratum"):

> Because of Bracketing Erasure at the end of each level only the structure assigned at a given single level will ordinarily be available for reanalysis to operate on. This blocks cases like (28) [see below, RS] from being derived. The marked cases such

---

5. Mohanan (1982) has an equivalent Opacity Principle. See Chapter 4 for a discussion of these principles.

> as ungrammaticality are then distinguished as exceptions to Bracketing Erasure. If the bracketing assigned to grammaticality at Level 1 is retained into Level 2, then the above principles [see (27) above, RS] permit ungrammaticality.

The cases prevented from being reanalysed which Kiparsky cites are as follows:

(29)

*unequalize	*symphony orchestrate
*non-ficitionalize	*Music Departmental
*chair-personify	*outboard motorize
*sheet-metallic	*freak accidental
*witch doctoral	*white elephantine

We will return to these cases below. For now, note that Kiparsky's analysis predicts that non-rebracketing, as in (29), should be the normal state of affairs insofar as such words are not marked exceptional to the otherwise general principle of Bracketing Erasure. This prediction alone seems suspect; we have seen that Bracketing Paradoxes occur productively in many languages, including Warlpiri, English and Russian and it seems therefore unlikely that one can get away with saying that they are a marked occurence.

There are some much more important problems with Kiparsky's analysis, however, which make it untenable as a solution to the problem, and which also suggest that the solution must abandon the assumption that there is only one level of representation for word-structure. The first problem is that appealing to exceptional marking of certain words explains nothing about why some words behave in this way and others do not. What precisely

is it about _-ity_ nouns that makes them exceptions to Bracketing Erasure, whereas, for example, verbs ending in _-ize_ are not? I shall suggest below that the relevant difference is syntactic productivity and semantic compositionality, but given this, we immediately suspect that the solution has nothing to do with the phonologically motivated Bracketing Erasure Convention in the first place. Kiparsky does not propose a reason for the difference in behavior of certain affixation processes so we really cannot say what might be going on[6].

The second problem is already suggested by the first. As noted, the Bracketing Erasure Convention has generally been assumed within theories of Lexical Phonology and Morphology to account for the general failure of phonological rules to look inside morphologically complex words and see the morpheme boundaries which were present at earlier levels. Note that this is a different requirement from the Strict Cycle since the latter condition does not preclude including a left or right bracket in the statement of a phonological rule and having that rule apply on a cycle on which the relevant bracket is as deeply embedded as you like, so long as reference is

---

6. Hargus (1985) argues that in Sekani, an Athapaskan language, there is evidence that certain affixes are exceptions to Bracketing Erasure on phonological grounds, and she concludes that there is a principled motivation behind Kiparsky's proposal that affixes may be exceptional with respect to this principle. However, her evidence is irrelevant for Bracketing Paradoxes of the kind discussed by Kiparsky insofar as the latter have no accompanying phonological evidence for being exceptions to Bracketing Erasure; see below.

made to some material introduced on the current cycle. The Bracketing Erasure Convention, however, does preclude such a rule application for the trivial reason that internal brackets will never be present at subsequent strata. Claiming that words like grammaticality are exceptions to Bracketing Erasure means, however that such rule applications ought to be allowed for such forms. That is, not only should such words be exceptional in that they allow reanalysis to take place, but they are in principle exceptional phonologically. Yet there seems to be no evidence of this expected exceptionality. This is in one sense a weak argument against Kiparsky's proposal since one could merely claim that, for English at least, no relevant phonological rules exist which would show up the exceptionality which the lack of Bracketing Erasure would allow. In another sense, however, it is a serious objection since such rules could in principle exist and we would thus have a weaker theory of morphology insofar as more derivation types would be allowed, which would otherwise be ruled out. The fact of the matter is that, as we shall see, the Bracketing Paradox cases are not the right type of data to force us to abandon the Bracketing Erasure Convention[7].

---

7. As it turns out, Bracketing Erasure may well be ill-motivated in Lexical Phonology and so may well not be a principle of the grammar anyway; again, I discuss this point in Chapter 4. Nevertheless, a problem still remains for Kiparsky insofar as his version of Lexical Phonology assumes this convention and the appeal to its exceptionability on the basis of Bracketing Paradoxes seems at best ad hoc. In any event, if the BEC is thrown out altogether as I suspect it should be, we then have the reverse

None of this is to say that evidence could not be found, perhaps in some language other than English, that Bracketing Paradox cases are exceptional phonologically in the relevant way; i.e., that there is phonological evidence for exceptional Bracket Retention in these cases and these alone. Nor is there any principled reason why a (phonologically based?) motivation could not be given for the exceptionality of certain forms and not of others. Failing such at the present time, however, I turn to a discussion of a theory of morphology which allows that words have more than one level of representation.

1.1.10 Pesetsky again.

The remarkable thing about every analysis of Bracketing Paradoxes investigated so far is the implicit assumption that there is one level of representation for morphological forms and that the Paradoxes must be somehow resolved at that level. But why make that assumption? One might imagine that words are represented in one way or another at different levels of the grammar--e.g., at D-structure, S-structure, PF and LF. Why should we take it as the unmarked assumption that there is only one level of representation for words? Rather, what seems natural is that there should be different word structures at the different levels, each

---

problem, for Kiparsky's theory, of no longer being able to rule out the non-existent "blends" whose non-existence Kiparsky attributes to the BEC.

determined by principles at those various levels. This is the thesis which I shall be pursuing in this work, and it is also the thesis which Pesetsky introduces in his quite novel--and already quite influential--analysis of the Bracketing Paradox problem.

An important point which an analysis such as Kiparsky's obscures is that in every case of Bracketing Paradox which we have examined so far the two bracketings which seem required to coexist with one another were required for two different sets of reasons; that is, a morpho/syntactic set of requirements seemed to require the one bracketing, while a phonological set of requirements seemed to require the other. To take the familiar ungrammaticality example again, the bracketing [[un grammatical] ity] is required because the prefix un- needs to be adjacent to an adjective; that is, it syntactically subcategorizes for an adjective. On the other hand, the bracketing [un [grammatical ity]] is required for reasons having to do with, e.g., the theory of Stratum Ordered Phonology. Pesetsky makes the quite sensible suggestion that these two types of requirements, namely syntactic requirements such as c(ategorial) selection and phonological requirements such as Stratum Ordering are properties of different components of the grammar. So, Pesetsky suggests that words have representations at more than one level; and these representations need not be isomorphic to one another.

The question, then, is what are the two grammatical levels of

representation relevant for representing the two levels of word-structure? Pesetsky suggests that they are S-structure for the phonologically motivated bracketings and LF for the syntactic/semantically motivated bracketings. This decision is by no means arbitrary. Following the analysis of Russian quantification presented in his dissertation (Pesetsky, 1982), he claims that categorial selection takes place at LF, rather than at some other syntactic level. In this sense, then the ascription to LF of the syntactically motivated bracketing of words is a sound one. The phonologically motivated bracketing exists at S-structure; from there it can be fed directly into PF to be interpreted with a bracketing isomorphic to its S-structure bracketing by the phonological component.

Given that there are two levels of representation for words, we need now to define a mapping between those two levels. Pesetsky suggests that precisely the right mapping already exists, and that it is the rule of QR originally discussed in May's dissertation (1977). His analysis of a case like ungrammaticality, therefore, proceeds as follows. At S-structure we have the following structure, which, as already noted, is precisely the structure motivated on phonological grounds:

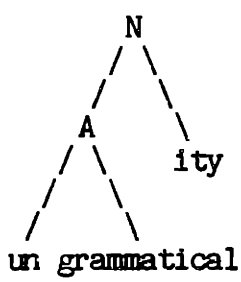
(30)



No categorial information is represented here since it is not relevant.

Now, taking, as Pesetsky suggests, QR as being a rule which applies generally in the mapping from S-structure to LF, and in particular is not restricted to apply to bona fide quantifiers (of which -ity is surely not one), we can assume that the mapping from the structure in (30) to the desired LF structure is mediated by this rule. Two possibilities exist, each of which Pesetsky presents. In the first possibility, QR does not leave a trace and in the second it does. The first possibility is the simplest, and yields a structure like the following:

(31)



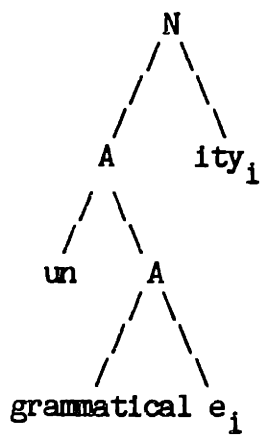
That is, assuming the percolation conventions of Lieber (1980) (see Chapter II for the exact statement of the conventions), the node dominating ungrammatical will be labeled as an adjective. This will mean that the LF representation of ungrammaticality will have -ity sister to an adjective and hence satisfied as far as its c-selection requirements are concerned. The node dominating the whole word will, again by the same conventions, be labeled N.

The other possibility, namely that affix QR does leave a trace, is somewhat more complicated, and requires the following stipulation (= Pesetsky's 73):

(32) Trace Stipulation: The trace of an affix belongs to the null category class 0.

The representation of the word ungrammaticality under this analysis will be as follows:

(33)



Given that the trace e is of category 0, and again given Lieber's percolation conventions, the node dominating grammatical followed by the trace of -ity will have its features percolated from the adjective and will hence be an adjective. We will return to the issue of traces in morphology below.

Pesetsky, with the mechanisms he introduces, is able to account for a number of Bracketing Paradox examples. There is a problem, however, in

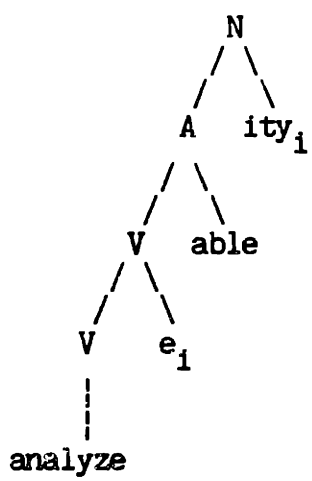
that QR as the mapping between S-structure and LF generally allows for reordering in interpretation. Famous examples such as the following serve to illustrate this point:

(34)
a. Every man loves a woman.
b. (∀x): x a man [(∃y): y a woman, x loves y]
c. (∃y): y a woman [(∀x): x a man x loves y]

In the (b) interpretation of (a) has [every man] and [a woman] in the order of their appearance in S-structure. (c), however, has them reversed in intepretation.

Given Pesetsky's analysis, what is to rule out words like *analyzitiable, which is obviously bad, and which could have the LF representation as follows (assuming the trace analysis though this is irrelevant for the present purposes)?

(35)

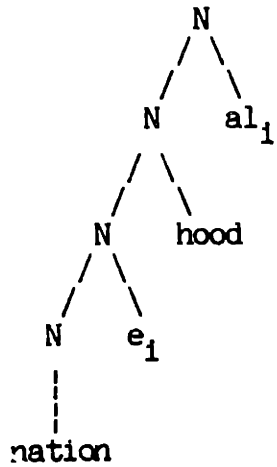


A quick examination of this construction reveals that it is well-formed

according to Pesetsky's principles, since _-able_ is sister to a verb, which means that it is happy, and _-ity_ is next to an adjective, which means that it should be happy. Pesetsky, however, does have a way of ruling out these constructions. If we suppose that the mapping to LF from S-structure occurs cyclically, and if constructions must be well-formed at every level then we can claim that *_analyzitable_ is ruled out because at the level at which _-ity_ is attached to the verb _analyze_, its subcategorization requirements ([+A, +V]) cannot be satisfied. So *_analyzitable_ is out for the same reason as *_analyzity_ is out.

The same logic does not carry over to examples of the following form, first pointed out in Sproat (1984a) and also discussed by Pesetsky in the most recent incarnation of this paper. Take a word such as *_nationalhood_, which is obviously horrible. Now, there is no question that _national_ is well-formed, and given that, there is no reason why one could not then submit it for further affixation. Affixing _-hood_ to it of course violates the subcategorization requirements of that affix, which takes a noun and forms nouns. But this ought to be salvageable by the same method used to save _ungrammaticality_, namely by raising the internal affix:

(36)



And the word could, in principle have a reasonable enough interpretation such as 'pertaining to nationhood'. This is perhaps a little unfair insofar as _-al_ may well not be semantically compositional, hence not fully productive. As I shall discuss in the context of my own analysis of Bracketing Paradoxes, even classic Bracketing Paradox reanalyses seem to fail consistently enough with this affix (c.f., Kiparsky's *_witch-doctoral_ example given above.) Even so, one gets the impression that not only do such forms as *_nationalhood_ not get the required interpretation, but that they would be horrendously bad under any circumstances. There are other examples which ought to be allowed to reorder in interpretation, and a list is given below (from Sproat, 1984):

(37)
*adventurouslet, *codifylet, *resistanceing, *activateish

There interpretation would be as follows:

(38)
as of a little adventure (adventure+let+ous), to make (like) a little code (code+let+ify), the action of being resisting (resist+ing+ance), to make somewhat active (active+ish+ate)

Now, the same criticism about semantic opacity and productivity could be raised against any of these examples, but I believe that there is a strong sense in which these forms are out for much more basic reasons than this. Actually at least one of the reorderings is an almost plausible word, namely ?codeletify; the interpretation of this word is clear enough, but it is equally clear that *codifylet absolutely cannot have this interpretation, and indeed is bad altogether

Pesetsky suggests that the solution to this problem, if it is real, is in a restriction of the following form:

(39) String Vacuousness Restriction
The terminal string output of QR does not differ from the terminal string input.

We return shortly to this proposed condition. Pesetsky notes, however, that there are potential counterexamples to (39), cases where we might want to say that reordering in interpretation as predicted by the QR model of rebracketing does occur. One of these forms is misreattach and other verbs of that form. Pesetsky claims that this can be interpreted with re- outside misattach. He gives the following example:

(40)
<Bill misattached the antenna once.> Then, after checking the instruction, he went back and mis-re-attached it!

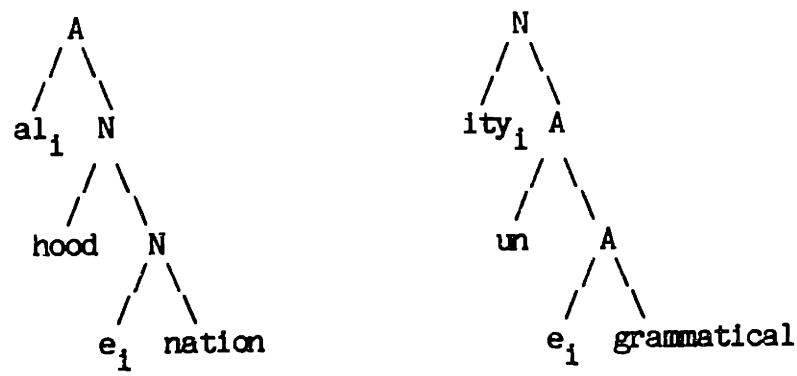
According to Pesetsky, this sentence is marginally allowable with the obviously intended interpretation that misreattached means 'misattached for a second time.' However Pesetsky also notes that many speakers do not share this intuition. I fall into the latter category.

His other example comes from Spanish. According to Pesetsky, there are dialects of Spanish where esferoidita 'sphere+sort-of+diminutive' is ambiguous between its straightforward reading of 'a little spheroid' and the QR reanalyzed reading of 'almost (sort of) a little sphere'. The analogous thing can also happen with esferitoide, whose most natural interpretation is 'almost (sort of) a little sphere' and reanalyzed interpretation is 'a little spheroid'. Nevertheless, Pesetsky notes that these reanalyzed interpretations are at best marginal.

So in evidence against the String Vacuousness Condition we have a couple of quite marginal examples, and in evidence for it we have some fairly clear cases which do not allow reordering in interpretation. Needless to say, then, the evidence is strongly in favor of having such a condition. Yet this Condition is by no means unproblematic since in order for it to make any useful claim whatsoever about the putative morphological mapping between S-structure and LF it necessitates the assumption, otherwise unmotivated I believe, that linear order, as opposed merely to scopal (i.e., c-command) relations, is relevant for LF representations. To see why this is so imagine rather that LF representations, as is usually

assumed, care only about c-command and not about linear ordering. Then the representation of ungrammaticality and *nationalhood will be as follows, where, linear order is irrelevant:

(41)



In these representations there is no apparent difference between *nationalhood and ungrammaticality. That is, there is no sense in which -ity is 'adjacent' to grammatical in a way in which -al is not 'adjacent' to nation. Of course, there is a difference, and that is that in the ungrammatical case we are dealing with two suffixes, and in the grammatical case we are dealing with a suffix and a prefix. But these facts, which are based on linear ordering, should in principle be irrelevant at LF. This is not to say that a mapping principle could not be found; some modified version of the Mapping Principle which I will be proposing (section 1.2.2.) could be reinterpreted as characteristic of the mapping from S-structure to LF rather than of the mapping from S-structure to PF as I shall suggest. But it is also clear that Pesetsky's String Vacuousness Restriction fails unless we make otherwise unwarranted stipulations about LF; note again that

it is never necessary to assume that syntactic QR is restricted by linear ordering.

There is one other class of cases which Pesetsky argues his analysis can handle and those are cases where idiosyncratic interpretation is involved. An example is the word rarity which can refer to the state of being rare or to a particular kind of object which is rare. The contrast is brought out by the following examples (modeled on Pesetsky's):

(42)
a. We were shocked by the rarity of these books.
b. A first edition of The Canterbury Tales is a rarity.

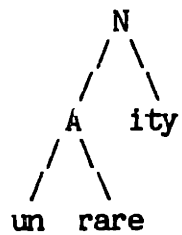
In the second sentence, rarity has the idiosyncratic interpretation. Now, Pesetsky notes that un- cannot prefix to the idiosyncratic word rarity. So there is no counterpart to (42b):

(43)
a. We were shocked by the unrarity of these books.
b. *A first edition of The Canterbury Tales is an unrarity.

In the case that QR does not leave traces the analysis is straightforward if we make the quite reasonable assumption that the idiosyncratic interpretation is only available when the items which are to be idiosyncratically interpreted are sisters, thus forming a constituent at the relevant level of interpretation. This characteristic of idioms has been noticed elsewhere insofar as idiosyncratically interpreted forms tend to act like frozen expressions, resisting passivization and wh-movement. In any event, given this assumption, along with the assumption that

morphological QR does not leave a trace, the LF representation of unrarity must be:

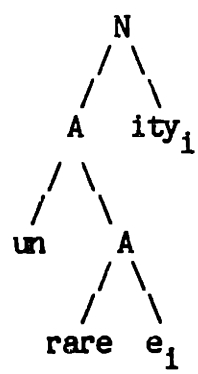
(44)



In this case, then, -ity and rare are not sisters, do not form a constituent, and are therefore not interpretable idiosyncratically.

The trace analysis is somewhat more problematic since it would yield an LF structure like the following:

(45)



But one may well ask why the trace cannot stand in for the raised affix in interpretation; that is, why rare + e cannot be interpreted as if the -ity were still there for the purposes of the semantics. As Pesetsky suggests, this is a troubling, though perhaps not shocking problem. He notes for

instance that there are cases of syntactic movement which preclude idiomatic intepretation for instance Heavy NP Shift:

(46)
a. ?John gave Bill e yesterday the answer.
b. *John gave Bill e yesterday the finger.

On the other hand, there are also cases of syntactic movement which allow for idiomatic interpretation. So, with sentence idioms such as 'The shit hit the fan' or 'The cat is out of the bag' it is perfectly possible to raise the subject of the idiomatic sentence:

(47)
a. The shit seems e to have hit the fan.
b. The cat seems e to be out of the bag.

(48)
a. The shit was believed e to have hit the fan.
b. The cat was believed e to be out of the bag[8].

I have no trouble interpreting these sentences idiomatically, and that suggests that traces may, in some cases, be interpreted in place of idiomatic antecedents. So the issue is not at all clear, and there may be a genuine problem for the analysis of morphological QR with traces which Pesetsky presents (see my own account, section 1.3.1.2.)

Pesetsky also analyses the behavior of thematic role assignment in

---

8. Noam Chomsky has suggested to me that the contrast between (46) and (47-48) might have to do with the difference between internal arguments and external arguments. We will not pursue this question here.

deverbal forms. I shall not discuss his examples here, but rather return to the problem of deverbal morphology in Chapter II, where I present my own analysis of the data.

1.1.11 A Summary So Far

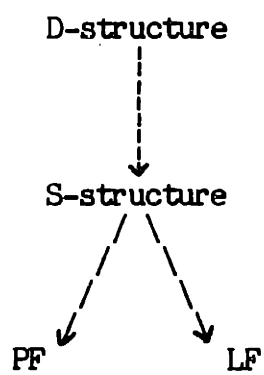
I have traced the history of the Bracketing Paradox problem from the time when such forms were first noticed in the literature to the present. Various analyses have been presented and many of them have proved unsatisfactory for various reasons. Pesetsky's analysis, which in many ways is closest to the one I wish to propose, is quite problematic since it turns out to be hard to find a motivated constraint for the putative application of QR to morphological forms. I should note that I am not claiming and neither am I convinced that genuine QR does not apply to affixes in some cases, especially affixes, such as the _-er_ comparative, which may well be quantifiers after all (see Section 1.3.1.1. for a discussion of this point.) However, I would claim that it is simply wrong to characterize Bracketing Paradoxes as a product of the mapping from S-structure to LF. I therefore turn now to my own analysis, and the arguments in support of it. But first some background on the model of grammar I shall be assuming.

1.2 The Mapping from Syntax to PF.

1.2.1 Background: The model of grammar.

The general model of grammar which I shall be assuming is as outlined in Chomsky (1981), developed in subsequent work within the Government and Binding framework and diagrammed in (49) below:

(49)



The model consists of two syntactic levels, D-structure and S-structure. D(eep)-structure, which is intended to be a pure representation of the thematic structure of lexical items in that all and only the arguments required by a particular item are present at that level. In particular, expletives such as it and there will be absent from D-structure representation. The mapping between D-structure and S(hallow)-structure is implemented by the general rule Move-alpha (or more generally

"Affect-alpha", Lasnik and Saito (1984)), which can "move anything anywhere", and which is itself restricted in its applications by general conventions such as the Projection Principle, Bounding Theory, Case Theory, Binding Theory, and so on, which either govern the rule's application or else apply to its output. I assume familiarity with these principles and will therefore abstain from a description of them.

From S-structure, there is a mapping to both of the so-called interpretive components, L(ogical) F(orm) and P(honetic) F(orm). It is generally assumed that the logical interpretation, along with the interface to the semantics resides at LF, whereas in the PF component lies the phonological part of the grammar, though what exactly counts as "phonological" has rarely been clearly stated. The mapping to LF, called Quantifier Raising, or QR was first studied in detail by May (1977) and has received a substantial amount of attention since that time. The corresponding mapping from S-structure to PF has received far less attention than it deserves; it is my intention here to take a shot at the problem.

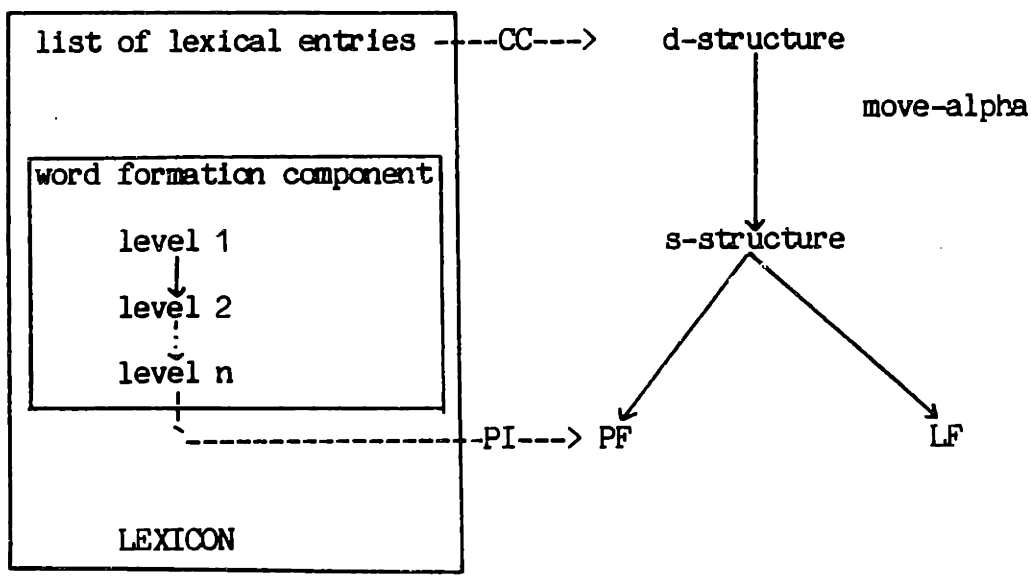
One may well ask if there is anything missing from the diagram above. In particular, where is the component of the grammar which has generally been called the lexicon? Perhaps the most explicit and comprehensible model of grammar, within the GB-framework, which attempts to give a precise location of the lexicon with respect to the other components is that

developed in Pranka (1983). One of the great insights, I feel, of this work was the realization that the process that had traditionally been termed "lexical insertion" was in fact the conflation of two separate processes which should be viewed as residing in two separate components of the grammar. The first of these processes Pranka called "categorial construction" (CC); this takes place at D-structure and consists in the "projection of the lexicon of the categorial information contained in the lexical items that constitute a sentence" (Pranka, p. 15). In her model, this information consists of nothing more than feature bundles. In many current views of grammar, following Stowell (1981), there has been a sharp turn away from the use of phrase structure rules, a strong preference rather being shown for deriving phrase structure entirely from the lexical properties of particular lexical items. Thus in a sentence like 'John hates rutabagas' there is no need to have a rule that rewrites a VP as a V followed by an NP since we need in any event to specify that hate is a verb which takes one internal argument, the 'hatee.' Other constraints on phrase structure can be seen to fall out from principles of Case-assignment, theta-assignment, and so on (see, Stowell, 1981; Koopman, 1984; Travis, 1984; Sproat, 1985a). Pranka's CC therefore is simply the projection at D-structure of the properties of lexical items.

The other half of lexical insertion is what Pranka terms Phonological Insertion or PI. This occurs, somewhat unsurprisingly, at PF, and consists in the insertion into the structure for a derived sentence of the

phonological material associated with particular lexical items. Pranka does assume that there is a separate (Stratum-Ordered) morphological component and that the relationship between the material inserted by CC and that inserted by PI is governed by the workings of the morphology. Her model is diagrammed below:

(50)



Pranka also assumes an operation, much in the spirit of Marantz (1981, 1984) which allows syntactic operations to merge feature bundles originally inserted by CC into D-structure; this operation takes place between D-structure and S-structure. Thus feature bundles at S-structure may contain more information than those at D-structure, accounting for the fact that in many languages single words may contain elements which one would want to argue are syntactically separate at D-structure. At PF, PI maps the appropriate phonological forms to the feature bundles present at

S-structure. The lexicon, than, is the machine which generates phonological forms which will meet up with their syntacticaly derived counterparts at PF. This idea is similar to one developed by Marantz (1981) which suggests that lexical insertion occurs at every level and that whenever affixation occurs by some syntactic operation of merger, the lexicon is called upon to provide an appropriate form to be inserted into syntactic structure.

There are a couple of conceptual problems with Pranka's model to which we now turn. First of all, it is not clear what she views as being the relevant list of "lexical entries". Is it just morphemes, or might it contain fully formed morphologically complex words? The former conception is surely wrong simply because many morphemes simply do not have the right properties to be inserted into D-structure. For example, if we take a derived nominal such as destruction what we want to say is that what "projects from the lexicon" is the nominal properties of the word, plus whatever properties it has inherited by being derived from destroy. But there is no sense in which the two lexical items destroy and -ion project independently of one another; I hope to clarify the problem of projection in the next chapter. So what we want to say is that morphologically complex words can project their feature bundles. But if this is the case, it is not obvious what the function of the Stratum-ordered morphology is in this model. If morphologically complex words are in principle available for CC, then why do we need a machine to generate such words? In the

normal conception of Lexical Morphology (c.f., Kiparsky, 1983a; Mohanan, 1983), the Stratum-ordered lexicon is the very machine which derives the morpho-phonological and morpho-syntactic properties of words. Yet some of those properties must already be present before we enter the "lexicon," according to Pranka's model. Perhaps, then, what she intends is that the Stratum-ordered machinery is only present for purposes of generating the phonological form of words--i.e., that it essentially determines the phonological well-formedness of the words to be inserted by PI. If so, it is hard to see why this isn't merely a part of the PF component, rather than being a separate component of the grammar; after all, if the PF component is an interpretive component which concerns itself with the well-formedness of phonological expressions then we ought to expect principles which determine such well-formedness to apply at that level. So, I shall argue in this thesis (see Chapter 4), the Stratum-ordered phonology, insofar as it is correct, does nothing more or less than determine phonological well-formedness at PF.

The other problem with Pranka's model has to do with some recent work by Mark Baker (1985a) having to do with a principle he terms "The Mirror Principle". Baker notes that languages which have complex morphology marking syntactic operations exhibit in that morphological marking the same order as one would argue the corresponding syntactic operations occur in. For example, take the following examples from Quechua (due to Muysken (1981)):

(51)
a) Maqa-naku- ya- chi- n
beat-recip-dur-cause-3s
'He is causing them$_i$ to beat each other$_i$.'

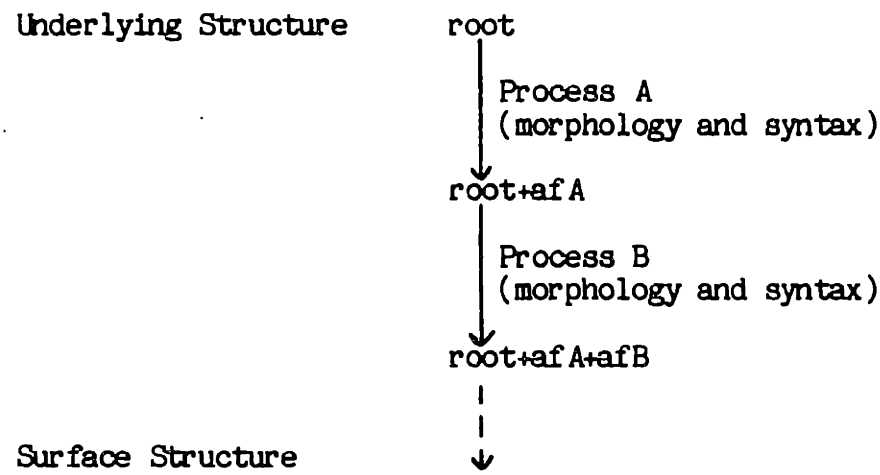
b) Maqa-chi- naku- rka-n
beat-cause-recip-pl- 3s
'They$_i$ let someone$_j$ beat each other$_i$.'

Baker observes that the difference in morphological structure in these two examples is also indicative of a difference in syntactic structure. In the first example the reciprocal affix is internal morphologically to the causative affix. Now, the interpretation, namely that the causee is identical in reference to the reciprocal, the object of 'beat', is consistent with a (D-structure) syntactic structure where the causee and the reciprocal are clausemates and where the subject 'he' is represented as the subject of the abstract verb CAUSE, which heads a matrix clause embedding the 'beat' clause underneath it. Causative formation thus occurs, in effect, after the interpretation of the reciprocal and the subject of 'beat' as coreferential. In the second sentence, the causative affix occurs internal to the reciprocal, suggesting that causativization occurs before the interpretation of the reciprocal and the subject of CAUSE as coreferential. This hand-in-handness of morphology and syntax, Baker states as follows:

(52) Mirror Principle
Morphological derivations must directly reflect syntactic derivations (and vice versa).

He suggests a model of grammar which incorporates this principle by having a constant interaction between morphology and syntax in the mapping from D-structure (which, again, is intended to be a pure representation of the lexical properties of lexical items, and which therefore is devoid of the kinds of morphological intricacies discussed by Baker) to S-structure:

(53)



Thus, syntactic operations, along with the morphological operations which spell them out go hand in hand.

Returning to Pranka's model, it is hard to see how one can derive this effect (except of course by stipulation) if the morphology is a separate "black box", operating wholly under its own principles and separate from the lexicon. That is, if the syntax derives feature bundles which are associated with phonological representations at PF, what principle determines that the phonological representations should preserve the order in which the feature bundles were built up, given that the lexicon derived

the phonological representation according to its own principles.

So, Baker's findings suggest that there is a very intimate relationship indeed between syntactic operations and the morphology which spells out those operations. Also, given the validity of my other criticism of Pranka's model, we bring into question the status of the lexicon as a word-building device. One of the questions which Pesetsky raised in his paper (1985) is the status of the component called the lexicon. "Is there a lexicon?", he asks, and I would like to suggest that the answer is no. That is, there is no lexicon if "lexicon" is taken to be a word-formation machine, a component of the grammar completely separate from syntax (and possibly from phonology). What I shall suggest, rather, is that the lexicon is what has come to be termed the "permanent" lexicon; i.e., a list of morphemes, and idiosyncratic words, phrases, etc., with their properties. General principles applying at the various levels of the grammar such as D-structure, S-structure, PF and LF will determine the well-formedness of various collocations of these lexical items. The subsequent chapters of this thesis will be an investigation into precisely what the principles which apply at the various levels, in particular in the Syntax and the PF component, might look like.

Notice that I am not claiming that there are no principles of word-structure; in Chapter 2 I will be investigating the properties of deverbal nouns, and it will turn out that at least some of the syntactic

principles which determine well-formedness of such collocations at the X^0 level will be only partly similar in their behavior to the behavior of principles which determine well-formedness of projected (i.e., X^n, $n>0$) items. If one insists that, for instance, the set of principles which applies to collocations which would generally be called lexical in fact must be termed "the lexicon", I have no particular objection, but notice that this is only an artificial distinction: special principles presumably apply to noun phrase structure, yet no one, I suspect, would be willing to argue that there is a separate "Noun-Phrase Component" in the grammar. Why should it be any different for words? Well-formedness for word structure (as originally suggested, I believe, by Lieber) can be derived in many cases from principles of categorial seletion, and as I shall also argue, thematic role assignment and possibly Case assignment (see also Fabb, 1984, on this last possibility), principles which may operate <u>differently</u> at the word level, but which are syntactic principles nonetheless. On the phonological side, there are principles of lexical stress assignment as well as principles of phrasal stress assignment. Lexical Phonology, will consist of phonological rules and principles which apply differently within phonological words only insofar as phonological words are generally built up by the affixation of items which phonologically "select" for other items. Post-lexical phonology will consist of phonological operations which do not occur under affixation. This will hopefully become clearer in Chapter 4, where I discuss this topic, but I note here that more or less

the same suggestion has actually already been made within Lexical Phonology, by Mohanan (1982) insofar as he suggests that in many cases the same rule may apply both lexically and postlexically, the differing application resulting from the properties of the lexicon in the former case and the lack of those properties in the latter.

I shall suggest, then, that all words (whether they be "derived pre-syntactically") or built up by the operation of syntax, a la Baker, have a representation at a level of syntax. Derived words, such as derived nominals in English, for instance, will be represented as bracketings of morphemes at D-structure and will be left unaffected in the mapping from D-structure to S-structure. Their well-formedness or ill-formedness will be determined by syntactic principles such as categorial selection, theta role assignment, and so forth. Syntactically determined morphology will be derived via syntactic affixation (following, of course, the Mirror Principle). I will have very little to say about this kind of morphology henceforth[9]. The syntactic structures will then be mapped onto phonological representations, which will themselves be subject to rules and principles of the phonology. To take a simple example, the word grammaticality will have two syntactically represented morphemes, which we shall represent as GRAMMATICAL and ITY. GRAMMATICAL is an adjective, and

---

9. Note that syntactic morphology of this kind may well involve Bracketing Paradoxes. For such a case from Georgian, see Marantz (1984b).

we shall notate this by subscripting it with A, as per normal usage. ITY is an affix which subcategorizes for adjectives and forms nouns (we will state the affixation requirements of ITY more fully below). This we abbreviate with the subscript <A,N>; in general, in such a pair <A,B>, will be the property of a morpheme which selects for another morpheme of class A, and produces one of class B. This notation is really just a variant of the notation suggested in Lieber (1980). The syntactic representation for the word grammaticality, then, will be as below:

(54)

$$[\ \text{GRAMMATICAL}_{A}\ \text{ITY}_{\langle A,N\rangle}\ ]_{N}$$

This, I would claim, is the bracketing for this word at both D-structure and S-structure, and presumably also LF. It is obviously well-formed according to the principles of categorial selection. In addition, its semantic interpretation will be derived from the particular semantic entries for the two affixes, and the general principles of compositional semantics. One important thing to note about this structure, and to bear in mind during the rest of the discussion in this chapter is that there is no linear ordering implied in the diagram above. In fact, I claim that linear ordering is irrelevant, though not necessarily non-existent for syntactic structure, at least that of words, (see also Marantz, 1984c). The relative order of morphemes, we may assume, is due to the phonological properties of those morphemes, such as the property of being a prefix, or a suffix, and so forth. Therefore, in no case should linear ordering in a

particular syntactic representation be construed as being important[10].

At PF the phonological information about the word is represented. So, the fact that -ity is a suffix, and that it attaches to latinate stems, the fact that grammatical is a stem/word which is latinate, and so forth are all represented here. So, the phonological representation for (54) will be:

(55) $[[\text{gramatikal}]_{[+\text{latinate}]}\text{ity}]$

The question which now arises is: What governs the mapping between the two levels? I turn now to this question.

1.2.2 The Condition on Mapping Between S-structure and PF

Marantz (1984c) has suggested that there is an adjacency relation on mapping between morphological structure and phonological structure. His statement of the mapping principle is as follows:

(56) Mapping Principle
If [A] is sister to [B] in morphological structure then [A] must concatenate with [B] in phonological structure:

1. left/right concatenation: last member of [A] adjoins to

---

10. I skirt the issue of whether some linear ordering, e.g. phrasal ordering, is relevant at S-structure, as argued by Travis (1984), Koopman (1984), and Goodall (1984). It will make no difference to my analysis either way.

first member of [B] (left concatenation) or first member of adjoins last member of [B] (right concatenation). A "member" could be the whole unit, the left/right-most constituent, or anything in between in size.

2. nonlinear concatenation (e.g., Arabic verbs)

The basic idea, then, is that if A and B are sisters in morphological (i.e., syntactic) structure, they are adjoined or adjacent in phonological structure. That this is the right kind of principle will become abundantly clear in the discussion below.

However, although this is certainly the right idea, there is a technical problem with Marantz's statement of the Mapping Principle, at least if the above is to be taken literally as a formal definition of the Principle. Take the standard example ungrammaticality, which Marantz's Principle is intended to cover. We want the Mapping Principle to explain why a syntactic representation like [[un grammatical] ity] can be represented at PF as [un [grammatical ity]]. There is no problem with -ity: since that is a sister to ungrammatical at S-structure, it can adjoin to the last member of that collocation, which is precisely what it does, since it affixes to grammatical. However, un- would appear to violate Marantz's condition: it is a sister to grammatical at S-structure, hence it should adjoin to grammatical at PF. But it does not; it adjoins to grammaticality, which is the container of grammatical, a relation which technically ought to preclude this from being a proper adjunction for this particular S-structure representation; that is, the left/rightmost member of a

morpheme can be as large as the morpheme itself, but not larger. I shall remedy this situation in the definition of the Mapping Principle to be given here[11].

First of all, however, let us define some basic notions which will be useful in giving the formal definition of the Mapping Principle. We can imagine a lexical entry to be a two-place vector, the first entry of which is the syntactic representation of the entry, and the second of which is the phonological representation. We will adopt the convention of naming such a vector by a primed upper-case spelling of the normal phonological or orthographic representation of the morpheme, and the syntactic entry by the corresponding unprimed upper-case spelling. For example, the following are lexical entries of English:

(57)

a. GRAMMATICAL' = $\langle GRAMMATICAL_{A}$, gramatikal$\rangle$

b. UN' = $\langle UN_{\langle A,O\rangle}$, /ʌn-$\rangle$

c. ITY' = $\langle ITY_{\langle A,N\rangle}$, -ity$\rangle$

(The notation <A,O> on UN merely means that that morpheme affixes to an adjective and that it does not change the category of the item to which it attaches. In Lieber's (1980) theory, this would be stated by saying that UN' does not have a category of its own.) We also define the following

---

11. This deficiency, as pointed out to me by Marantz, is not characteristic of the approach taken in Marantz (1984a), where concatenation and adjunction are separate processes (see Marantz, 1984a, pp. 286-9).

useful relations and predicates:

(58)
a. $S(M)=_{def}$ the syntactic entry for **M**.
b. $P(M)=_{def}$ the phonological entry for **M**.
c. **sis**(**A B**), a predicate which applies in the syntactic representation and which is true iff **A** and **B** are sisters
d. **suf**(**A**), a predicate which applies at PF and which is true of **A** iff **A** is a suffix.
e. **pre**(**A**), a predicate which applies at PF and which is true of **A** iff **A** is a prefix.

We will now introduce the operator *, which can be read "is adjacent to," and which applies to pairs of phonological representations, and produces a third phonological representation which is the result of adjoining the two original representations according to the principles of autosegmental phonology in the case of affixation, or simply concatenation in the case of sisters which do not actually affix to each other[12]. This operator is commutative so that **A*B=B*A**. It is not associative, so that **(A*B)*C** is not generally equal to **A*(B*C)**.

However, it will generally be the case, in concatenative morphology (though not in non-concatenative morphology, which is why we need *), that

---

12. Thus * (and also ^) are deliberately ambiguous between operators, which is what they are mathematically, and predicates which is what they are interpreted as linguistically. For the purposes of this word it will suffice to interpret an assertion such as (**a*b**) as being a shorthand for the following: This phonological representation is such that **a** is adjacent to **b**.

either A or B is specified as a suffix or a prefix so that they adjoin in a particular order. We therefore introduce ^, an ordered (non-commutative) operator defined as follows:

(59) Definition of ^ for Affixation:
If
(i) suf(B), then A*B=A^B≠B^A;
(ii) pre(B), then A*B=B^A≠A^B.

In fact, it will more generally be the case that ordering is specified for syntactic reasons too; e.g. the language may be head-final or head-initial, thus ordering the constituents at PF; in this case we would introduce ^ to replace *. Although ^ is not commutative, it is associative, unlike *, in that if (A^B)^C counts as a phonological representation for some X, then so do A^(B^C) and (A^B^C).

We now define the relation PH which relates syntactic representations to phonological representations as follows:

(60) The Mapping Principle.

(i) If B is a morpheme, then PH(S(B))=P(B).
(ii) If sis(A B) then PH(A B)=(PH(A)*PH(B))[13].

This is more or less equivalent to Marantz's principle of adjacency, but as

---

13. In English this says the following: (i) if B is a morpheme then the phonological mapping of the syntactic representation of B is just its phonological entry, i.e. PH(B)); (ii) if A and B are sisters, then PH(A) and PH(B) are adjacent; i.e., the phonological representation of [AB] is such that PH(A) and PH(B) are adjacent.

we shall see, it does not suffer from the technical problem discussed above.

To see how this all works, consider the simple case of the word grammaticality. The necessary stipulations are listed below:

(61)
a.) GRAMMATICAL'=<$GRAMMATICAL_A$, gramatikal>
b.) ITY'=<$ITY_{<A,N>}$, ity>
c.) **suf**(ity)

As we have seen, the syntactic representation of the word is [GRAMMATICAL ITY]. Now, given the mapping principle, this will convert as follows: Since **sis**(GRAMMATICAL ITY), **PH**(GRAMMATICAL ITY)=(**PH**(GRAMMATICAL)***PH**(ITY))=(gramatikal*ity) and since **suf**(ity), this is equivalent to (gramatikal^ity), which is the result we want of course.

At this juncture, it will be informative to discuss how suppletion fits into this model. How do we represent the fact that went stands in for GO and PAST, which would normally be expressed in the PF component as two morphemes, but for the fact that go is irregular? I suggest that this is a fact about PF, and that there is, in such a case, a conversion rule which maps phonological adjunctions onto single phonological morphemes. So, for instance, in English there is an entry for went as follows:

(62)
went=go^ed

I assume that this rule is subject to the Elsewhere Condition (Kiparsky, 1983a) in the sense that whenever such a rule exists, it will supersede the more general output of the concatenation of two morphemes.

There also is a kind of inverse of suppletion, which is common in morphological systems and which we shall term "antisuppletion." This is where two or more phonological morphemes represent a single syntactic morpheme. An example from English would be the word produce, where the PF representation consists in the concatenation of pro- and -duce, and where we might assume there is no internal representation as far as the syntax is concerned, for the simple reason that the two morphemes of the phonological representation act merely as phonological formatives, having no productive syntactic[14] or semantic contribution to the words which they form. So PRODUCE' will have an entry as follows:

(63) PRODUCE'=<PRODUCE, (pro, duk)>

The following statements about the phonological morphemes will be necessary:

(64) **pre**(pro), **suf**(duk)

With all of this preliminary material out of the way, then, we turn to a

---

14. Though one could argue that since Xduce forms are invariably verbs, then -duce must have a syntactic representation and percolate its V feature to the word that dominates it.

discussion of Bracketing Paradoxes.

1.2.3 Simple Cases of Bracketing Paradoxes.

Let us now turn to an example like ungrammaticality. We list the necessary facts below, some of which are already familiar:

(65)
a. $UN' = \langle UN_{\langle A,\ 0\rangle}, \Lambda n\rangle$
b. $GRAMMATICAL' = \langle GRAMMATICAL_A, gramatikal\rangle$
c. $ITY' = \langle ITY_{\langle A,\ N\rangle}, ity\rangle$
d. **pre**(Λn)
e. **suf**(ity)

It will turn out that there are somewhat more stringent requirements on the affixation of ITY' than suggested here, but the definition above will serve the purpose for this discussion. As we have seen, the syntactic requirements dictate the following syntactic bracketing for ungrammaticality

(66) $[[UN_{\langle A,\ 0\rangle}\ GRAMMATICAL_A]_A\ ITY_{\langle A,\ N\rangle}]_N$

Now, the question which arises is whether [Λn[gramatikality]] is an acceptable phonological bracketing for ungrammaticality; it will turn out from the definition of the mapping relation given in the last section that it is. It is easy to prove this:

Since **sis**(UN GRAMMATICAL),
PH(UN GRAMMATICAL)=(**PH**(GRAMMATICAL)***PH**(UN))=(gramatikal*Λn). Since

pre(/ʌn), this is convertible to (/ʌn^gramatikal). Now, since sis([UN GRAMMATICAL] ITY), PH([UN GRAMMATICAL] ITY])=(PH[UN GRAMMATICAL]*PH(ITY)). But this is just ((/ʌn^gramatikal)*ity) which is convertible to ((/ʌn^gramatikal)^ity). But since ^ is associative, this is equivalent to (/ʌn^(gramatikal^ity)), which is the desired bracketing.

So, we have shown that the desired rebracketing in the PF component is permissible, given the Mapping Principle introduced in the last section. But what forces this rebracketing to occur in the first place? The answer, of course, as already discussed in previous sections, is that phonological principles applying in the PF component decide which of the permissible bracketings is the correct one In this case, the fact that, say, ity is a Stratum I suffix whereas /ʌn is a Stratum II prefix will force the bracketing [/ʌn [gramatikal ity].

It is fairly obvious that any of the bracketing paradoxes introduced in the literature such as Pesetsky's (1979) original examples from Russian, and other examples such as unhappier can be handled analogously to ungrammaticality. Just to give another couple of simple examples, consider first of all the English verb forwent, past tense of forgo. Now, it seems that we want to say that the syntactic bracketing is [[FOR GO] PAST], for the simple reason that forgo is a verb in its own right, and in particular does not have a meaning which is compositionally derivable from for and go. PAST, we may assume, is attached outside at S-structure as a result of

attaching INFL to the outside of V. However, we want to say that the phonological consituency is [for went] for the obvious reason that went is a single phonological constituent, suppleting for GO and PAST. That this mapping is allowed is clear, since [[FOR GO] PAST] will convert into ((for^go)^ed), which is equivalent to (for^(go^ed)), and given the suppletion given in (62) in the last section, this will convert to (for^went).

Note also that this Mapping Principle will allow rebracketing of structures with arbitrary depths of embedding. Take the example ungrammaticalification (discussed also in Sproat (1985b)) which is admittedly contrived, but which I believe is interpretable. The word is in fact ambiguous between "the action of doing the reverse of making grammatical," which has the bracketing [[UN [GRAMMATICAL IFY]] ATION], where un- here is the reversative verbal prefix, and "the action of making not grammatical", which would have the structure [[[UN GRAMMATICAL] IFY] ATION], with the un- being the adjectival negative prefix. Now, in both cases the PF bracketing must be [ʌn [[gramatikal] ifi:k] a:tiʌn], since ʌn- is Class II, and all of the suffixes are Class I. But this result follows from the Mapping Principle; take for instance the case where UN is the adjectival prefix, since this involves the greatest amount of rebracketing. This will map to (((ʌn^gramatikal)^ifi:k)^a:tiʌn), which, given that ^ is associative, can rebracket--and, indeed, must do so since ʌn- must be external to the suffixes--to (ʌn^((gramatikal)^ifi:k)^a:tiʌn).

So, we see how the Mapping Principle developed here can handle arbitrarily deep rebracketings, something which is apparently necessary.

1.3 Handling Bracketing Paradoxes.

We now turn from these rather trivial examples to an examination of harder cases, some of which have appeared in the literature already and have been reviewed above. First of all, we examine two points raised by Pesetsky (1985).

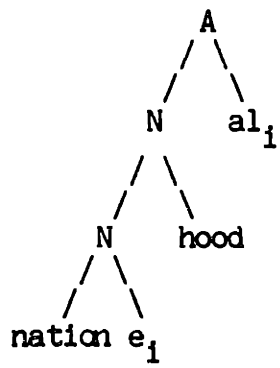
1.3.1 Handling Pesetsky's data

The two facts I wish to discuss here are the behavior of idioms with respect to rebracketing and the non-existence of reordering in reanalysis. We examine the second point first.

1.3.1.1 No reordering in analysis

Let us consider again the example *nationalhood, which we pointed out above, would be allowed to have the following LF structure under Pesetsky's analysis, assuming that there is no "String Vacuousness Condition":

(67)



Now, under the analysis of word-structure presented here, this "LF" representation for *nationalhood would have to be the S-structure representation for this word. That is, since the syntactic subcategorization requirements of the affixes must be met in the syntax, the only way for this group of morphemes to correctly bracket is as follows:

(68) $[[\text{nation}_N\ \text{hood}_{\langle N,\ N\rangle}]_N\ \text{al}_{\langle N,\ A\rangle}]_A$

Now, it is fairly clear that this could not map to [[nation al] hod] at PF since the Mapping Principle would yield ((nātiʌn^hōd)^al), and since ^ is not commutative there is no way to reorder these affixes to get the other bracketing. Of course, the bracketing that is actually yielded by the Mapping Principle would be ruled out since -al is Class I and -hood is Class II, or in Fabb's (1984) terms, -al affixes to [+latinate] stems, whereas -hood produces only [-latinate] stems.

What about the reputedly good cases of reordering that Pesetsky

discusses such as the misreanalyze case from English and the esferoidita example from Spanish? Obviously my analysis cannot handle these and insofar as these are quite marginal (I can at least say that with certainty with regard to the English examples), this is probably all for the better. Nevertheless, as I suggested in the previous discussion of Pesetsky's analysis, I have no evidence that QR or some similar mapping never applies to affixes. Rather, I suggest that if it does, it has nothing whatever to do with the large class of examples called Bracketing Paradoxes. It is interesting too that in such cases, at least some of the affixes involved might actually be quantifiers, hence making it plausible that something like QR might be involved after all; so mis- is a negative prefix, and hence has some of the semantic properties of a quantifier, and re- means roughly "again", which is also arguably quantificational in nature. As possible evidence of these assertions, I offer the following examples, using the syntactically separate words incorrectly and again:

(69)

a. Duane Gish incorrectly analyzed the transition from reptiles to mammals; he completely ignored the existence of Diarthrognathus. After so doing, Gish **incorrectly analyzed the reptile/mammal transition again.**

b. Duane Gish analyzed the reptile/mammal transition again. But as it happens, he **incorrectly analyzed the reptile/mammal transition again**; he failed to note the existence of Diarthrognathus.

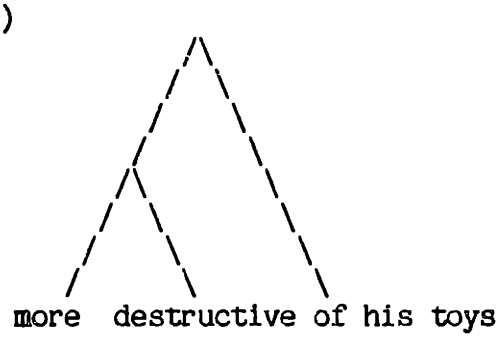
In the first case the meaning of the boldface sentence is that Gish misanalyzed the relevant data for a second time. In the second example, which has the same linear ordering as the first, the meaning is that he

incorrectly reanalyzed the data. It may be, of course, that the difference in meaning here is ascribable to a difference in syntactic structure, in which case the examples are irrelevant for arguing for the quantifier-hood of items meaning again and incorrectly. On the other hand, it may be that the difference lies in differing LF representations, paraphrased as follows:

(70)
(a) For some E, again(E), where incorrect(E), E an event of analyzing the reptile/mammal transition.
(b) For some E, incorrect(E), where again(E), E an event of analyzing the reptile/mammal transition.

There is one other example that has been discussed by Pesetsky and others and which might arguably involve a quantificational affix, and that is the comparative affix -er which, of course, participates in Bracketing Paradoxes. It may even be plausible to suggest that it raises. Fabb (1984: pp. 117-121) even goes so far as to suggest that -er raises to affix to the entire AP at LF. This is also true of the degree modifier more. Take, for example, the phrase more destructive of his toys. This, according to Fabb, must have an S-structure such as

(71)

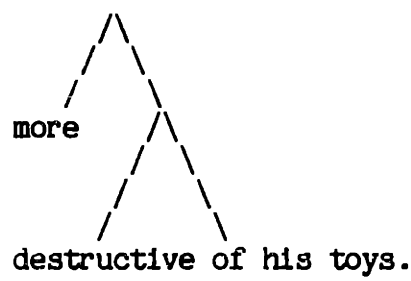


for the simple reason that the phrase more destructive can be topicalized out, stranding the complement, thus suggesting that more destructive is a constituent independently of the complement:

(72) More destructive though John is [__of his toys],....

However, this is not the LF representation according to Fabb; rather he suggests that it is as follows:

(73)



The reason is that the interpretation of this phrase can be "he was in a state of being 'destructive of his toys'; now he is more so." And as Fabb observes:

> The phrase does not mean that he was in a state of being destructive, now he is more destructive, and he is more destructive specifically of his toys.

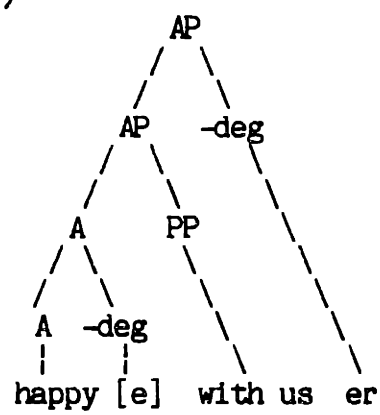
Be that as it may, the alternative LF bracketing does seem possible, as in:

(74)
Mary is far more destructive of her toys than she is of her clothes.

What this clearly means is that Mary is in a state of being more

destructive with respect to her toys than she is with respect to her clothes. This suggests the opposite LF bracketing from Fabb's example, in fact, the S-structure bracketing suggested above. Notice that if we were to say that Fabb's bracketing were right for this example, then following his reasoning, we would have to say that the sentence means that Mary is in a state of being destructive of her toys and being destructive of her clothes, but that she is more [destructive of her toys] than [destructive of her clothes]. Since the LF placement of more would be outside the AP, then we are essentially claiming (as Fabb implicitly does for the 'more destructive of his toys' example) that more is comparing two APs, which refer to different states of affairs; the state of affairs of being destructive-of-her-toys and the state of affairs of being destructive-of-her-clothes. These are essentially just two different predicates, and we would be thus comparing two utterly different states of affairs. But this does not appear to be what's going on here, since the semantics of the construction are clear. To reiterate, the meaning would appear to be that Mary is more destructive vis-a-vis her toys than she is vis-a-vis her clothes. So, we have suggested that more need not attach to AP, though it may do so. What about -er? Fabb suggests that the LF structure of a phrase 'happier with us' is as follows:

(75)



So the sentence John is happier with us now can mean that John was in a state of being happy-with-us and now he is in a state of being more happy-with-us. On the other hand, a sentence like Mary is happier with her toys than she is with her new clothes suggests the opposite bracketing, with the unraised degree modifier, as argued above. So there does not seem to be anything to force the LF raising, though it seems as though it may occur.

Still, although it may not be clear that -er must raise, there is nevertheless a genuine problem here. What forces degree modifiers such as -er to occur outside un-? Under the LF analysis of rebracketing, this must happen otherwise the interpretation of un-X-er as 'more not X' and not 'not more X' cannot be forced. Under the analysis of Bracketing Paradoxes presented here, assuming that the possibly real LF raising of degree modifiers has nothing to do with the rebracketing, then the question comes down to how we force the syntactic bracketing [[un X] er], and do not allow

the alternative bracketing. I suspect that the answer may have to do with the syntactic subcategorization requirements of the prefix un- rather than with any properties of -er. Although we stated above that un- attaches freely to adjectives, this is not quite true and there are restrictions. Aronoff (1976) shows that un- attaches more or less productively to adjectives ending with certain suffixes (p. 63):

(76)
a. $X_V$en (where en is the marker for past participle)
   c.f.: uneaten, unbought, unkilled, unbroken....
b. $X_V$ing
   c.f.: unsatisfying, uninspiring, uninteresting....
c. $X_V$able
   c.f.: unparsable, uncomputable, uninsertable, underivable....
d. Xy
   c.f.: unworthy, unhappy....
e. Xly
   c.f.: unseemly, unsurly....
f. Xal
   c.f.: unconditional, ungrammatical uncomical....
g. Xful:
   c.f.: uncareful, unmindful, unsorrowful....
h. Xlike:
   c.f.: unwarlike, un-MIT-like, undoglike....

Examples exist, to be sure, where un-X forms where the X is not an adjective ending with one of the above suffixes: unfit, unkind... But these seem to be less common; c.f., the oddness of ?unnice, ?unodd, ?unsad..., though some of these, such as unsad may have to do with Allen's (1978) restriction on attaching un- to negative bases, discussed above (section 1.1.3.). In any event, it seems plausible to suggest that un-X forms are either listed, as in the case of unfit and unkind, or select for

adjectives headed by particular affixes. This property of head selection is, of course, nothing new in morphology (see, also, the discussion of Guerssel's theory above). many cases abound, one of the more famous ones being the selection of _-ment_ for verbs formed with the prefix _en-_: _enjoyment_, _enrichment_, _employment_, _encasement_.... If this is the right analysis of _un-_, this will force the syntactic representation to have _un-_ inside _-er_ for the simple reason that _un-_ does not select for _-er_ as one of the adjectival heads to which it attaches. The representation for _unhappier_ will thus be:

(77)

[[UN [HAPP Y]] ER]
[+un]

1.3.1.2 Idiosyncratically Interpreted Words.

We now turn to the discussion of the issue of lexical idioms, a topic which will also be important in our answer to some of Kiparsky's examples of non-rebracketing. Take, again, Pesetsky's example _unrarity_, where, as discussed above, it is impossible to interpret _rarity_ as having the idiosyncratic interpretation of 'a particular item which is a rare object of its type.' The reader will recall that this was fairly naturally accommodated in Pesetsky's analysis with "QR" leaving no traces (and somewhat less naturally if "QR" does leave a trace) under the assumption that idiomatic interpretation of a pair of items requires that those items

be sisters--i.e., part of the same constituent. We can carry the same explanation readily over into our analysis here: that is, the reason that unrarity cannot have the idiomatic interpretation is that in order for rarity to be interpreted idiosyncratically, it will be necessary for the syntactic intepretation of the word to have the following bracketing:

(78) [UN [RARE ITY]]

But this structure will obviously be ruled out because of the subcategorization requirements of UN.

1.3.2 Kiparsky's Examples

We now turn to a discussion of the examples discussed in Kiparsky (1983c) for which he notes that rebracketing is impossible. I shall suggest that in many or all of these cases, the reason rebracketing is impossible is because the particular Level I suffix is simply either not productive syntactically, or not predictable in its semantic contribution to the word it is forming. This will mean, then, that words having that suffix will merely have to be listed, and will thus behave precisely like lexical idioms (such as rarity) with respect to Bracketing Paradoxes.

We give his examples again here (see section 1.1.9.):

(79)

*unequalize	*symphony orchestrate
*non-fictionalize	*Music Departmental
*chairpersonify	*outboard motorize
*sheet metallic	*freak accidental

*witch doctoral *white elephantine

Let us start with the easy examples. First of all, there is no reason to suppose that elephantine is anything but a listed form. The fact is that -ine is not a productive affix, although it seems to make a fairly predictable semantic contribution to the word to which it attaches. (Webster's dictionary lists it as having the meaning of "having the nature of, like". So elephantine means "like an elephant"; similarly adamantine means "like adamant" (diamond) and crystalline means "like a crystal".) But it is by no means productive. So, the following words do not exist:

(80)
*elementine "like an element"
*miss(i)line "like a missle"
*prostitutine "like a prostitute"
*nippopotamine "like a hippopotamus"

It seems reasonable, therefore, to suppose that the reason we do not get *white elephantine although we do have white elephant and elephantine is because the syntactic representation of elephantine must have ELEPHANT and INE adjacent; such words are therefore much like syntactic idioms in that they must be listed. Therefore the bracketing of white elephantine would have to be [WHITE [ELEPHANT INE]], which we would not expect to have any meaning associated with white elephant; note that white elephant is itself an idiom[15].

---

15. One interesting point is that phonological reanalysis in general is impossible when the left member is an idiom. Note the following:

Notice that what is being claimed is that the non-productivity of affixes such as _-ine_ will force the listing of the syntactic representation of all words in which they occur. They thus have the status of syntactic idioms; notice that _kick the bucket_ and _shoot the bull_ will have to be listed simply because they are semantically non-compositional. Similarly, _-ine_ does not _predictably_ attach to nouns, and hence all its formations are listed. This is in marked contrast to, say, agentive _-er_ (on which see Chapter 2), with which forms can be productively produced, hence generally eliminating the need for listing.

A similar point can be made for _-ify_ as in *_chairpersonify_. This suffix attaches to nouns and adjectives to make verbs, but it is not productive. When it does attach often it has a causative reading, though this is often vague as in the word _personify_, where there is only a remote relationship to _person_. Other ones are listed below:

(81)
liquefy "make liquid"
dignify "imbue with dignity"
glorify "imbue with glory"

---

*trigger-happier
*trigger-happiness
*uneasier (i.e, 'more uneasy' where _uneasy_ is not _not easy_).

This may indicate a stronger requirement for idioms than for other syntactic compositions, to the effect that the phonological representation of an idiom must have a bracketing isomorphic to its syntactic representation. And if this is correct _white elephantine_ might be out also because _-ine_ would have to attach phonologically outside _white elephant_.

putrefy "become putrid"

Again, however, it is probably the case that forms in -ify are generally listed as such, since other such forms are lacking:

(82)
*parsonify "make like a parson"
*squalefy "make squalid"
*lucefy "make lucid"

There probably are cases where -ify attaches outside Level II affixes, as in the somewhat contrived example ungrammaticalification, which I gave in a previous section; this is not a problem for my analysis, however, as these may be merely listed with the appropriate syntactic representation. The point is that we will not get productive rebracketings--such as chairpersonify, which we might otherwise expect given the existence of chairperson and personify.

A similar point can be made for *outboard motorize. -ize is a fairly common affix, but while it is fairly productive with adjectives, forming causatives, it has unpredictable meanings with nouns. Here is a sample of the meanings of X-ize nouns in relation to the meaning of the base:

(83)
motorize "provide with a motor"
hospitalize "put in hospital"
lionize "treat as a celebrity" (idiomatic)
crystallize "become a crystal"
winterize "prepare for winter"

It seems reasonable, then, to suppose that such -ize words are listed, and as such we would not expect rebracketing.

The same story will hold also of *symphony orchestrate. -ate is not a productive verb-forming suffix; it is not a predictable fact that, for instance, orchestrate means 'to arrange for orchestra'. Unequalize is interesting in that although it is odd, it is clearly better than the alternative with the Stratum I negative prefix in-, *inequalize, and hence may really be a case of rebracketing after all.

The examples in -al, namely *Music departmental and *freak accidental, and in *-ic, namely *sheet metallic, seem problematic insofar as they contrast with examples (also from Kiparsky) where reanalysis with these affixes is possible: arch-ducal, vice-presidential, cross-sectional, root-parasitic, set theoretic(al), lord mayoral, double helical. The situation with -al seems especially problematic insofar as it seems like a fairly productive affix, with a fairly predictable meaning of 'pertaining to':

(84)
presidential = 'pertaining to a president'
mayoral = 'pertaining to a mayor'
ducal = 'pertaining to a duke'

Somewhat less clear cases are accidental which does not really mean 'pertaining to an accident', but rather 'by accident' and doctoral, which means 'pertaining to a doctor' only in the restrictive sense where doctor is someone with a Ph.D. It may be, however, that we will have to say, in fact, that -al adjectives are listed, despite the apparent productivity of this affix: this may well be the reason why *freak accidental and *witch

doctoral do not work (but see footnote (14) above). This would mean too that cases where -al is apparently bracketed outside Stratum II affixes must also be listed. This is not as undesirable as it might seem, since even with particular bases which participate in such Bracketing Paradoxes, not all Stratum II prefixes or compounding with these bases yields a well-formed word when affixed with -al:

(85)
a) wind instrumental; but ??brass instrumental
b) cross-sectional; but *6.001 sectional
c) vice-presidential; but *club presidential

The same point can be made for the cases in -ic, such as root-parasitic. So while this word is relatively acceptable, it is probably listed nonetheless; compare ??dog-parasitic, ??lung-parasitic...

Note that certain affixes seem to be productive with semantically definable classes of words. For instance, -ist fairly productively attaches to words referring to fields of study, producing nouns which refer to persons who engage in such study (see Marchand, 1969, pp. 308-9; note, too, that there are problems with allomorphy in some of these examples):

(86)

physics	physicist
chemistry	chemist
botany	botanist
linguistics	linguist
phonology	phonologist

And these appear to productively form Bracketing Paradoxes:

(87)

atomic physics	atomic physicist
organic chemistry	organic chemist
cryptogamic botany	cryptogamic botanist
historical linguistics	historical linguist
metrical phonology	metrical phonologist

I have suggested, then, that the relevant criterion for whether or not, say, a Stratum I suffix will syntactically bracket itself outside a Stratum II prefix, or a compound, is the syntactic productivity or semantic transparency of that prefix. The fact of the matter is, however, that it is often difficult to assess the productivity of a particular affix, even when it is obviously found in a large portion of the vocabulary. What does it mean to say that an affix is productive anyway? I attempt to answer this question in the next section with a study of the affixation requirements of _-ity_, an affix which is seemingly productive in many classes of cases, and which also often occurs in Bracketing Paradoxes.

1.3.3 The Affixation of _-ity_

We start by examining a class of adjectives where the affixation of _-ity_ is completely predictable and regular. This is the class of deverbal adjectives in _-able_. A number of these are given below, and I believe that there is no question that the corresponding _-ity_ nouns are grammatical:

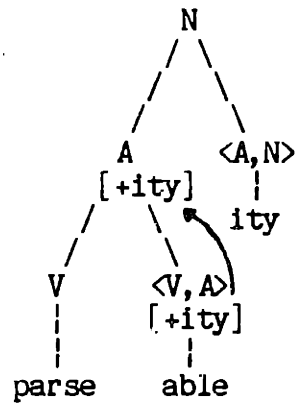
(88)

act	actable	actability
beat	beatable	beatability
catch	catchable	catchability
dent	dentable	dentability
eat	eatable	eatability

fry	fryable	fryability
get	gettable	gettability
.....		

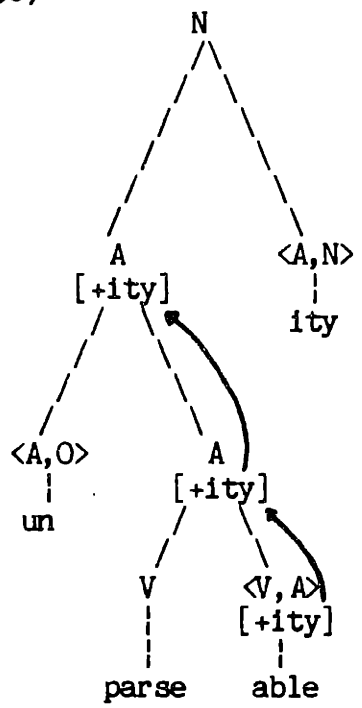
Many more such examples could be given. This suggests that there is a feature on -able which allows attachment of -ity to any words formed by that affix. Let us use an arbitrary diacritic, say [+ity]. A word like parsability will thus have the following structure:

(89)



According to the feature percolation conventions of Lieber (1980), the diacritic [+ity] percolates from the affix to the node dominating it. This is indicated in the diagram. Of course this feature will also percolate to higher A nodes if there are any and this predicts that un- ought to freely occur in un-X-able-ity forms:

(90)



This is correct, I believe:

(91)
unactability, unbeatability, uncatchability, undentability, uneatability, unfryability, ungettability....

Given this, we might expect to have to say that it is not sufficient for _-ity_ to attach to an adjective, but that in order for it to do so productively, there must also be an inducing diacritic[16]. What about affixes other than _-able_? _-Al_ seems to be productive enough: over an eighth of the _-ity_ words listed in the version of Webster's dictionary

---

16. See, again, Guerssel (1985) (and section 1.1.8. above), who makes the same point about _-able_ and _-ity_.

which is on-line on the TOPS-20 system are based on adjectives ending in al; this compares fairly favorably with -able adjectives, which form fully one third of the -ity words. Here are some of the al adjectives from that list:

(92) actuality, animality, bestiality, cardinality, centrality, eventuality, locality, morality, nationality, principality, reality, speciality, vitality

Nonetheless, it is also clear that many of these words have fairly idiosyncratic meanings which are not productive of the meaning of the base adjective plus -ity. So actuality does not primarily refer to 'the state of being actual,' but rather to a fact, something which is actually true; nationality does not mean 'the state of being national'; speciality is not 'the state of being special'; and principality is not the 'state of being principal.' Moreover, there are a number of -al adjectives which do not take -ity:

(93)
*accidentality, *doctorality, *ducality, *neurality...

This suggests, again, that despite the large numbers of -ality nouns, the suffix -al just does not allow of -ity affixation productively. This, then, is in marked contrast to -able which always allows -ity. So we probably need to say that -ality words are just listed. What then of cases like ungrammaticality, which have been the center of debate on Bracketing Paradoxes? If grammaticality just has to be listed as [GRAMMATICAL ITY] in syntactic structure then why is this particular Bracketing Paradox allowed

in the first place? The answer might be that ungrammaticality is also listed, again as [[UN GRAMMATICAL] ITY], and is not a productive product of the fact that -ity attaches freely to adjectives ending in a certain affixes. This may not be as undesirable a conclusion as it at first seems since naive speakers actually find words of this form odd, although they do accept un-X-ability words fairly freely. Nevertheless, it is still possible that in a subset of the -al adjectives, there is productive rebracketing going on, at least in the cases, like grammaticality~ungrammaticality, territoriality~extraterrioriality, metricality~extrametricality, where the -ity has the more or less productive meaning of 'state of'.

Adjectives ending in -ous are interesting. As discussed in Aronoff (1976), Xacious and Xocious regularly drop the -ous ending before -ity; c.f., mordacious~mordacity, precocious~precocity. Xecious adjectives regularly retain it, spelled as -os, though this is actually a misleading generalization since there are but a handful of -eciosity words (two in the TOPS-20 on-line Webster's); I suspect that Xecious does not form a natural class of adjectives. As Aronoff points out, other classes of -ous adjectives are unpredictable as to whether or not they retain -ous before -ity (Aronoff, p. 41):

(94)

nebulous	*nebulity	nebulosity
credulous	credulity	*credulosity

Un- does not seem to freely attach to -ity words derived from Xous adjectives:

*unperspicacity, *unprosperity, *unaudacity, *unscrupulosity, *unvivacity.

Many of these cases can probably be fairly well ascribed to the fact that un- doesn't attach to the base adjective either:

(95)
*unperspicacious, *unaudacious, *unvivacious.

Others, such as unprosperous and unscrupulous are okay, even though their -ity derivatives are not. I suggest, then that -ity forms of -ous adjectives are listed.

What we seem to have to say about -ity is that it attaches in the unmarked case to adjectives, and it forms nouns from them, but it may only freely do so if they are formed with a suffix, of which we have found exactly one so far, which "induces" the attachment. Again, this property of one affix inducing the attachment of another one is not unusual in morphology; c.f. the strong propensity of -ment to attach to words prefixed with en-, mentioned above.

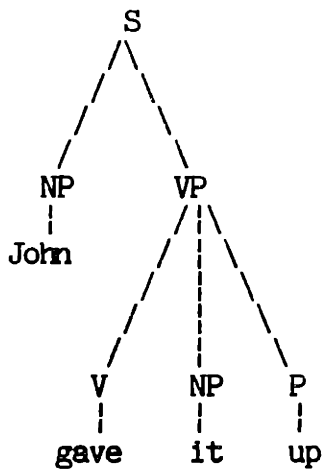
1.4 More on the Mapping Principle

1.4.1 A curious piece of English morphology.

We now turn to an interesting and informative problem in English morphology which turns out to have a fairly natural statement in terms of the Mapping Principle given here.

Verb-particle forms in English have inspired a fair amount of debate in the literature (see Yip 1978; Lindner 1981; Stowell, 1981 for discussions and analyses of various facets of this problem). One thing which generally seems to be agreed upon is that the syntactic representation of the verb particle constructions in sentences has the particle appearing essentially as a sister to the verb within the verb phrase. For example:

(96)

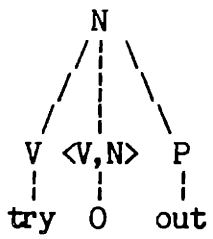


Suppose that we assume that as in the verb phrase the particle occurs as sister to the verb in <u>any</u> construction in which the verb could occur by

itself.

Now, when one of these verbs is nominalized via "null affixation", it will follow that the particle would become "frozen" in the sense that no nominal affixation could occur inside it and on the verb stem itself. That is, the structure will be something like:

(97)



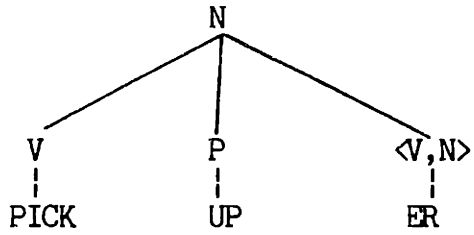
Nominal affixes, such as the plural, will be sister to the N node, and hence attach outside and could not attach to the verb stem (although it might be that they phonologically rebracket and attach to the particle). This prediction is confirmed by the following:

(98)

try outs	*tries out
make ups	*makes up
pick ups	*picks up
take outs	*takes out

What will happen when an affix such as the agentive <u>-er</u>, an affix whose syntactic and semantic properties I shall be investigating in the next chapter, attaches to the verb? The syntactic structure for such a form, given the assumption that <u>ER</u> must be sister to V and that the particle will also be sister to V is as follows:

(99)



This structure has the following properties:

(100)
1. **sis**(PICK UP)
2. **sis**(PICK ER)
3. **sis**(ER UP)

This immediately implies the following, given the Mapping Principle:

(101)
1. (pick*er)
2. (pick*up)
3. (er*up)

This means essentially that all of the morphemes are (non-fortuitously) adjacent, which is an impossibility. Something has to give. Plausibly, the part that gives is the statement (pick*up), which says that pick and up are adjacent; this is often observationally false at PF anyway, as in sentences with particle movement, for instance. We are thus left with:

(102)
1. (pick*er)
2. (up*er)

Now, since -er is a suffix, this converts to:

(103)
1. (pick^er)
2. (up^er)

But the only PF structure which allows for these conditions to be simultaneously met is one where _-er_ is essentially "reduplicated" as in _picker-upper_. Interestingly enough, such forms are quite productive, as pointed out by Yip (1978):

(104)

picker upper
putter downer
taker outer
maker upper
....

Other forms are possible, such as _picker up_ and _pick upper_, and these could be obtained by suppressing either one of the statements in (102).

This is obviously not the whole story; some forms are markedly worse with "reduplication": ??_looker arounder_, ??_fucker arounder_, *_screwer overer_... Presumably some phonological constraints do apply to such forms. But the crucial point is that there is no reason whatsoever to suppose that there are two _syntactic_ instances of _-er_ in any of these forms. The fact that two _phonological_ instances may occur is interesting, and follows with only a few auxiliary assumptions, from the Mapping Principle[17].

---

17. Ken Hale informs me that this kind of morpheme doubling is common in languages. See Appendix to this chapter for a discussion of a possibly related example in Berber.

1.4.2 Rebracketing of Separate Words

The next class of cases I shall examine are cases where there is a mismatch between the syntactic definition of a word and the lexical definition. A typical example of this phenomenon involves prepositions in many languages. The syntactic representation of a Prepositional Phrase presumably has a preposition sister to the NP which it governs, i.e.:

(105)
[P NP]

The preposition is a syntactically separate word. On the other hand, in many languages for a large class of prepositions, there is reason to believe that the phonological representation has the preposition bracketed with whatever the first member of the NP is, as a single word:

(106)
[[P X] ....]

Pranka (1983) has noted the existence of this phenomenon in several languages, and calls it "Fusion". One very clear case of fusion is the French word du which stands for the preposition de followed by the masculine definite article le. This is in marked contrast to the feminine definite article la which does not (obviously) fuse with the preposition. Now, presumably the syntactic structure of the PPs are identical in either case, namely:

(107)

$[_{PP}$ DE $[_{NP}$ LE ....]]

$[_{PP}$ DE $[_{NP}$ LA ....]]

The Mapping Principle will allow the preposition to concatenate with the first element of the NP, if we mark the preposition as being a prefix; this is probably a fairly reasonable analysis insofar as in many languages, prepositions also function as prefixes (c.f., Latin, Greek, English...), and they often behave identically to prefixes in their phonological behavior, as in the case of Welsh, to be discussed below. In the case of DE followed by LE, we will have the following suppletive rule:

(108) P(DE')^P(LE')---> du

Marantz has also discussed these forms (1984a), which he terms "Surface Merger" and relates them to the classic cases of Bracketing Paradoxes (1984c).

In the next subsection, I give a case study of one such system, the consonant mutating prepositions of Welsh. In the subsequent subsection, I discuss how my Mapping Principle might relate to proposals of Lobeck and Kaisse (1984) for handling similar data in other languages.

1.4.2.1 Consonant Mutation in Welsh

It is not my intention here to delve into the intricacies of the phonology of Celtic consonant mutation; for treatments of this, see

Thomas-Flinders (1981), Sproat (1982), Crill (1985), and somewhat less to the point Massam (1983), Lieber (1983) among others. I shall point out that however one treats the phonological changes induced under consonant mutation, we will need to assume that the phonology is lexical even when it appears to apply across word boundaries.

The major consonant alternations of Welsh are listed below:

(109)

Radical	Lenition	Spirantization	Nasalization
p	b	f	m̥ʰ
t	d	θ	n̥ʰ
k	g	x	ŋ̊ʰ
b	v	b	m
d	ð	d	n
g	(deletes)	g	ŋ
m	v	m	m
ɬ	l	ɬ	ɬ
r̥ʰ	r	r̥ʰ	r̥ʰ

Of all the mutations, lenition is by far the commonest. It occurs in the following environments, among others:

(i) After many prepositions: e.g., i 'to', kardiff 'Cardiff', i gardiff 'to Cardiff'

(ii) After some of the possessive pronouns: e.g. i 'his', ki 'dog', i gi 'his dog'

(iii) In feminine singular nouns after the definite article and in adjectives modifying feminine nouns: e.g. ər 'the', garð 'garden', mawr 'big', ər arð vawr 'the big garden.'

(iv) After the great majority of prefixes: e.g., kam- 'mis', deaɬ 'understand', kamðeaɬ 'misunderstand'.

(v) In compounds: kox 'red', di 'black', koxði 'brownish'.

Nazalization occurs in the following environments:

(i) After the preposition ən 'in'; e.g. kardið 'Cardiff', ən$^{h}_{J}$ardið 'in Cardiff'.

(ii) After the possessive və 'my'; e.g. ki 'dog', vən$^{h}_{J}$i 'my dog'

(iii) After prefixes such as an- 'not'; e.g. diolxgar 'thankful', aniolxgar 'unthankful'

Finally, spirantization occurs:

(i) After some prepositions; e.g. a 'with' kəɬeɬ 'knife', a xəɬeɬ 'with a knife'

(ii) After the possessive pronoun i 'her'; e.g. ki 'dog', i xi 'her dog'.

(iii) After some prefixes; e.g., tra- 'over', kevn 'back', traxevn

'again'.

These lists are by no means exhaustive, but they do give a fairly representative sampling of the environments in which the various mutations occur.

The question arises, given the wide range of environments listed above, exactly what component of the phonology the rules occur in. In fact, there is good reason to believe that the rules are lexical. Note first of all that they apply in "classical" lexical environments such as compounding and prefixation; lenition, which is particularly prevalent in the lexical phonology of Welsh, is also used to mark feminine gender on adjectives. And interestingly enough, the so-called syntactic applications of the rules fall into a very narrow class of cases, namely prepositions, and pronominal clitics, two classes which have been characterized in other languages as behaving phonologically as affixes. It is significant, for instance, that in no case in Welsh does a rule of mutation occur, say, between two nouns which just happen to be contiguous in the syntax.

Furthermore, there is evidence that mutation not only <u>can be</u> lexical, but in fact must be lexical; that is, it has properties which make it appear to be as lexical as any other arguably lexical rule. In particular, mutations are subject to lexical marking in that they are associated with particular prepositions. So Spirantization occurs with <u>a</u> 'with,' but not with <u>yn</u> 'in,' which induces Nasalization or <u>yn</u>, the particle which precedes

predicate nominals and adjectives, which causes Lenition. This is parallel to the clearly lexical cases of prefixes: so tra- 'over' induces Spirantization, whereas an- 'not' induces Nasalization and di- 'not' induces Lenition.

What then do we need to say about the prepositions and clitics which induce mutation? We will say that syntactically they behave as separate words but that phonologically they are marked as affixes. Assuming that phonological affixes in the unmarked case attach to the adjacent phonological word, then we will get the result that a preposition such as i 'to', for instance, will attach to whatever phonological word follows it. The lexical entry for this preposition will look, in part, as follows:

(110)
$I' = \langle I_p, i \rangle$
pre(i).

An NP such as i gardið 'to Cardiff', will have a synactic representation as follows:

(111)

$[_{PP}\ I\ [_{NP}\ CAERDYDD]]$

and a phonological representation as follows:

(112)
(i^gardið)

Similar examples can be drawn from Irish, as discussed by Pranka (1983) and also Massam (1983). In that language prepositions supplete with following definite articles, much like du in French. For example, the sequence le 'with' and an 'the' becomes leis an which, although it is written as two words, is really a single phonological word. The interesting thing here, discussed by Pranka, is that this suppletion can occur across sentence boundaries:

(113)

Níor fhan sí leis [an mbróg a thógáil]
not wait[+past] she with the shoe to pick-up
'She didn't wait to pick up the shoe.'

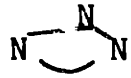
In this example, le syntactically governs the entire embedded clause [an mbróg a thógáil]. However, it phonologically fuses with the article of the object NP of the clause. This kind of behavior is expected given that the mapping from S-structure to PF only specifies that two items, such as le and an must be or can be adjacent, and in particular does not specify their PF bracketing. In the PF component, items which form completely separate syntactic words are thus free to form single phonological words. This is a Bracketing Paradox par excellence, going far beyond the domain of what one would normally think of as morphology. We now turn to a discussion of some proposals by Lobeck and Kaisse (1984) for handling phenomena similar to the ones discussed here and in Pranka (1983) and Marantz (1985a,b).

1.4.2.2 Other "Locality" Phenomena

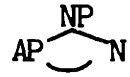
Lobeck and Kaisse (1984) discuss a number of cases of syntactically governed phonological phenomena. One such phenomenon is lenition in the Paleo-Siberian language Gilyak, the phonological side of which is a rule which spirantizes a consonant when following a vowel and voices it when following a nasal. This occurs in environments such as the following:

(114)

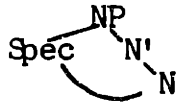
a. Noun-noun compound
   q'os 'neck'
   ŋə͡ẋos 'otter neck'



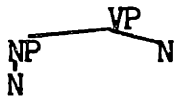
b. Adjective-noun
   təf 'house'
   pilar͡dəf 'big house'



c. possessive pronoun-noun
   pəx 'paint'
   ŋəŋ bəX 'our paint'



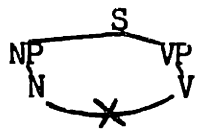
d. direct object-verb
   ŋas͡pəkzḍ 'throw away belt'
   ki͡vəkzḍ 'throw away shoe'



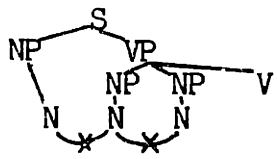
It fails to occur in the cases such as the following:

(115)

a. subject NP-verb
   taqo/pəkzḍ 'the knife disappeared'



b. NP subject-NP object; NP object-NP object.
ni/taqo/p'ə kən k'imd
I knife to-brother gave



Given the examples above, Lobeck and Kaisse are able to make the following generalization:

(116)
Lenition occurs between a and b where b c-commands a.

In discussing wanna contraction in English, they note that c-command between want and to is a requirement for this substitution to take place:

(117)
a. I want [to go home]
wanna

b. I [persuaded [the guest [you want]] [to come]]
*wanna

An example from Postal and Pullum shows that this condition is not sufficient, as Lobeck and Kaisse note:

(118)
I don't [want [[to flagellate oneself] to become standard
*wanna
practice in this monastery]]

In this case, although want does c-command to, it is separated from to by two S -nodes. Lobeck and Kaisse make the following generalization:

(119)
Contraction may occur when the host verb governs the minimal S' containing to.

In fact, since such examples involve infinitives, there is no reason to suppose that the embedded clauses are S', rather than S. Given this, another way of stating this generalization, and a way which fits the data as well as Lobeck and Kaisse's statement is the following:

(120)
Contraction may occur when the host verb is sister to the minimal S containing to.

In fact, assuming, as is generally done, that to is in INFL and INFL is the head of S, this reduces to:

(121)
Contraction may occur when the host verb is sister to S headed by to.

Looking back at the data from Gilyak, a similar generalization can be made:

(122)
Lenition occurs between a and b where b is sister to X headed by a.

A general condition on these syntactically governed sandhi processes then might be simply that one of the words whose phonological representation engages in the process is sister to the phrase headed by the other word. This implies, of course, given the Mapping Principle, that the first word and the phonological representation of the phrase must be adjacent. Given that the second word--the head of the phrase--is initial in the phrase and therefore adjacent at PF, then the rule can apply; the

rule is blocked from applying between items which are too far separated by the constraint that the item within the adjacent phrase must be the head.

1.5 On some cases which might seem to support morphological QR

In this last section I turn to some examples which have been cited in support of a QR model of morphological interpretation such as that outlined by Pesetsky (1985). In the first subsection I discuss data from Navajo as analyzed by Speas (1984). In the following subsection I speculate on some other cases which are similar to Navajo.

1.5.1 Morphological QR in Navajo?

One interesting thing about Navajo verbal morphology is that a good many verbs consist of a preverb and a verb stem where these two components do not occur separately, and thus are idiosyncratically interpreted as a single lexical item. One such verb is ya+ti' 'talk', where neither ya nor ti' appears separately. This is not odd in itself, but what is odd is that no matter how idiosyncratic the interpretation of the preverb plus the verb, the inflectional morphology attaches inside the preverb/verb complex; for example, yashti' 'I talk', where sh is the first person singular marker, and /l/, the voice marker, drops out by a regular phonological

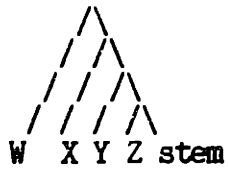
rule. In fact, quite a large number of affixes can occur between the preverb and the verb. And the interesting thing about these affixes is that many of them occur in the opposite order, relative to the verbal stem, to what is generally found across languages. The surface ordering is as follows:

(123)
preverb>ITER>DIST>DIRECT-OBJECT>MODE>SUBJECT>VOICE-ELEMENT>stem

Notice in particular that the object occurs outside the subject, which is counter to the ordering commonly found across languages with double agreement. (However, voice occurs within the subject and object markers which is common cross-linguistically.)

There are two problems here, then, namely how to get the preverb and the stem interpreted together, and how to get the affixes interpreted in the "right" order. Speas makes the suggestion that this is accomplished via a QR-mapping to LF. To see how this will work, let us imagine we have the following (S-structure) representation for a word, where W, X, Y and Z are all prefixes:

(124)

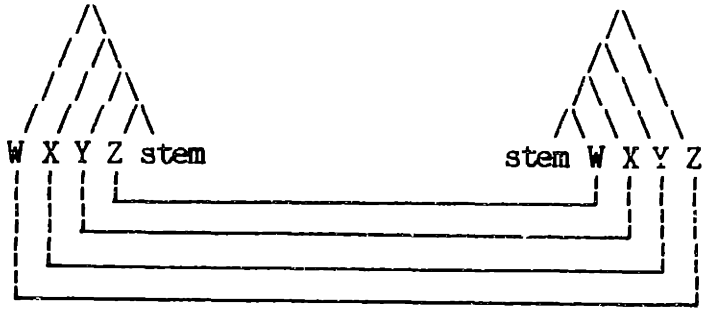


Speas makes the following claim:

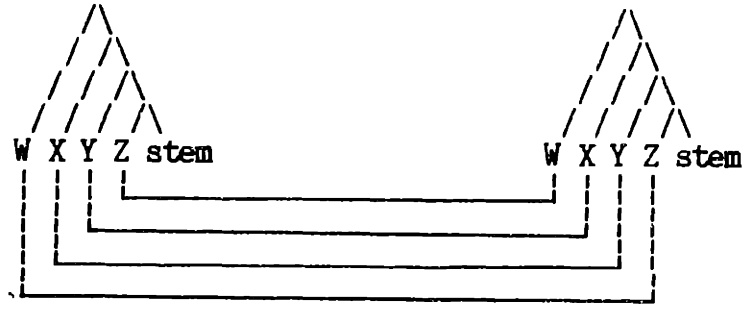
A common constraint on mapping from one level of representation to another is a "no crossing constraint." Insofar as such a constraint is empirically motivated in the mapping of syntax to LF, it is reasonable to expect that the mapping from the structure to the semantic representation of a word will also obey a "no crossing constraint." Assuming that dominance relations are relevant for LF but precedence relations are not, the prediction of a theory like that of Pesetsky et.al., supplemented by a no-crosssing constraint, is that surface poistion of affixes must be either precisely the order needed for interpretation of scope relations, or the mirror image of that order, as shown in (21) a and b, respectively. Any other order will violate the no-crossing constraint in mapping from word-structure to semantic representation, as shown by (21)c.

(21) [S-structure: RS] [LF: RS]

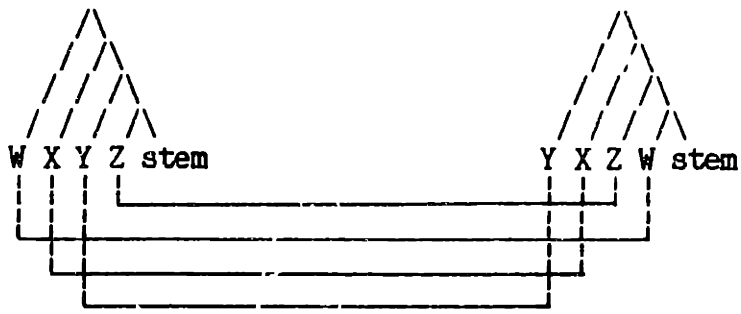
a)



b)



c) *



Actually, it is far from obvious that the "No-crossing constraint" is correct for QR, at least not the real QR that we find with quantifiers. Take a classic sort of ambiguous example like the following:

(125)
a. Two women saw three men.
b. (2 women)(x)[(3 men)(y): x saw y]
c. (3 men)(y)[(2 women)(x): x saw y]

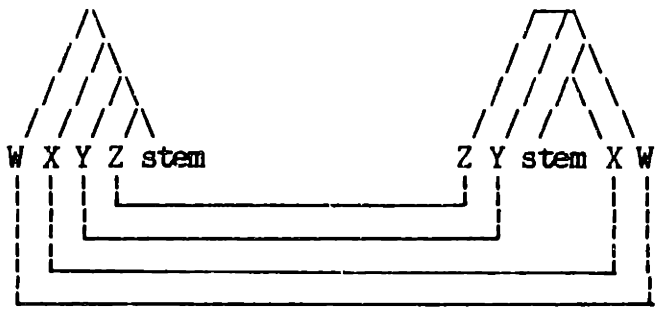
In either case we assume that QR has taken place, but only in the (c) reading is "No crossing" preserved.

Even granting that "No crossing" is a valid constraint on morphological "QR", Speas' claim—that Pesetsky's theory predicts that interpretation must occur either in the surface order or in the reverse of that order--is false. First of all, what forces affixes to raise? In the cases like ungrammaticality, one can certainly claim that -ity is raised so as to allow un- to be adjacent to an adjective, but it is not obvious what will force the raising of the Navajo affixes since they are all verbal affixes, and are certainly attached to a verb, even at S-structure. Yet if they need not raise, and supposing that only some of them do, then there is

nothing to prevent various "scrambled" orders of interpretation from occuring.

Say, then, that we do have a requirement that such affixes raise. Then, contrary to Speas's claim, if linear order is genuinely irrelevant for LF (as she rightly assumes), then only the reverse order of interpretation is possible, a result we surely do not want. Let us see why this is so. Speas claims that the linear order has been imposed in the diagrams above only for the purposes of presentation, but the fact is that they are crucial for allowing the S-structure order of interpretation to be maintained in the (a) example. It is only because the string W X Y Z is drawn on the right of the stem that the mapping works for this example; if you are unconvinced by this, try to come up with a diagram that represents this mapping (with W internal to X internal to Y internal to Z at LF) with the appropriate string of characters written on the left of the stem. The fact is, that the ordering in these diagrams is crucial, and, if it is to be claimed that these represent the legal mappings from S-structure to LF, then it is implied that ordering is relevant for LF too. But if such diagrams are representative of this mapping, then again, there is nothing to stop scrambled orders of interpretation. I have provided such an example in the diagram below:

(125)



If, on the other hand, linear ordering is genuinely irrelevant for LF, and assuming a No-crossing convention, we can easily see that only the reverse interpretation is possible, if affixes must raise. So, interpreting '>' as 'c-commands', and '<' as 'is c-commanded by', then we can see that if W>X>Y>Z at S-structure, then raising all the affixes should reverse that order to W<X<Y<Z.

So, we have seen that Speas' Navajo data do not provide the evidence for Pesetsky's rule of QR that they might initially seem to. They rather seem to cause problems for that analysis, since to claim anything interesting, we seem to have to stipulate a No-crossing convention which may not be otherwise motivated for the mapping from S-structure to LF, and possibly also assume that linear ordering is relevant (which, as we have seen, also leads to complications; we have actually already run up against this problem of linear ordering at LF in the discussion of Pesetsky's String Vacuousness Condition.)

As a solution to the problem, I propose the following. At S-structure

the verb and the preverb are represented as sisters. The affixes are adjoined to the verb, via syntactic affixation, in the "right" order of interpretation, so that the whole structure looks like the following:

(127)
[SUBJECT [MODE [DIRECT-OBJECT [DIST [ITER [VOICE [PREVERB STEM]]]]]]]

The Mapping Principle will necessitate that the following hold of PF:

(128)
(subject*(mode*(direct-object*(dist*(iter*(voice*(preverb*stem))))))

How then do we get the right surface form, with the affixes between the preverb and the stem, and linearly arranged with the most external affixes more adjacent to the stem? We seem to need to say that that each of the affixes is in fact an infix and that the most external infixes end up more internal towards the stem; the affixation would be non-concatenative. Interestingly enough, this is exactly what Speas claims is going on in the first half of her paper. She argues for a model of Navajo verb morphology where the affixes are infixed into phonological insertion frames. An example of a derivation is given below (= Speas' (17)):

(129)

hinishghal 'I arrived wriggling'

STEM STRATUM



stem conjugation

insert voice element



STRATUM I

no affixation

insert V in env. C_C

/i/



STRATUM II

insert /ni/ (= perfective mode) in env. CV__(C)C(V)C#

insert /sh/ (= subj. 1) in env. __CV(V)C#



sh+l+ghal--->shghal

OUTPUT hinishghal

So, both the perfective (= mode) and the subject are infixed into the stem. In fact, the ordering of infixation given in this example is voice before mode before subject (the exactly that necessitated by the syntactic

bracketing given above.) Given this, Navajo verbal morphology can thus be seen not to support Pesetsky's theory at all, but rather the Mapping Principle suggested here.

To see this in more detail, consider the syntactic representation of hinishghal, which we may take to be as follows:

(130)

```
[SH  [NI   [L     [H       GHAL]]]]]
 mode subj  voice  preverb stem
```

This will convert to the following, via the Mapping Principle:

(131) (sh*(ni*(l*(h*ghal)))))

Given that h is a prefix, this will further convert to:

(132) (sh*(ni*(l*(h^ghal)))))

Now, sh, ni and l are specified as infixes as part of their phonological subcategorization properties. So, we infix these affixes in the order specified by the output of the Mapping Principle, namely l first, followed by ni, then sh.

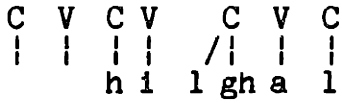
First l will insert between the preverb and the stem, to yield the following:

(133)

```
C  V  C   C  V  C
|  |  | / |  |  |
      h l gh a  l
```

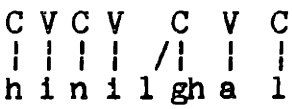
Assuming for the sake of convenience of exposition a Stratum Ordered phonology as envisioned by Speas, we will pass on to Stratum I, where /i/ will be inserted by phonological rule:

(134)



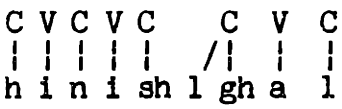
At Stratum II, ni will be inserted as specified by Speas in CV__CVC#:

(135)



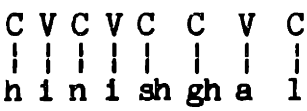
Finally, sh will be inserted in __CVC#:

(136)



We will then apply the rule converting the sequence /sh+l+ghal/ to shghal:

(137)



We have thus derived the surface phonological form of this Navajo verb from the representation derived via the Mapping Principle, which had the reverse scopal order. Morphological QR is thus superfluous.

Note, too, that the interpretation to * (adjacency) here is a perfectly coherent one; two objects may be said to be adjacent for the purposes of the phonology if the phonology combines those two objects via some process of affixation: i.e., either by simple concatenation (via the intermediate operator ^), or by infixation, or by more complex autosegmental means such as the non-concatenative morphology of Arabic (McCarthy, 1979).

1.5.2 Other Cases Like Navajo

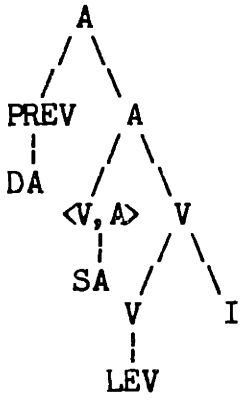
Of course, providing a solution to the Navajo situation along the lines proposed in the previoous section does not necessarily solve the problem of parallel constructions in other languages. It is not clear that Speas' infixation analysis for Navajo could be extended, for instance, to Georgian, which as discussed by Marantz (1984b), also has verbal forms of the following structure:

(138) [Preverb [Inflectional-Affixes Verb-stem]]

In Georgian, as in Russian or Navajo, the preverb is idiosyncratically associated with the verb stem. In fact, the situation in Georgian is even worse than in Navajo, since Georgian has, for example, adjectives which are derived from verbs where the adjective forming affix occurs between the preverb, if there is one, and the verb stem. So, from čer- 'write,' which has no preverb, we have sa-čer-i magida 'writing table,' with the adjective forming prefix sa- and the suffix -i (data from Aronson, 1982, p. 181). A

verb such as da-lev- 'drink,' which has the preverb da-, has the related form da-sa-lev-i 'beverage,' where the prefix sa- intervenes between da- and lev-. This would presumably have the following syntactic structure:

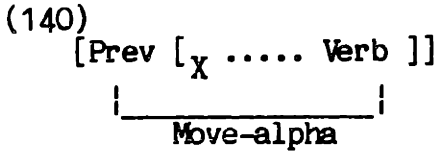
(139)



Such a structure, however, would preclude the analysis of such preverb situations, suggested in Sproat (1984a), whereby the reason that structures of the form [$_V$Prev[$_V$Inflection-Verb]] are allowed is that preverbs could attach to any verbal projection of their selected verbal stems; in this case, DA is attaching to an adjective.

The suggestion of Marantz (1984b) is more to the point, namely that in these situations a Move-alpha relation holds between the morphosyntactic representation of such forms and their phonological representation. In particular, the following relation holds:

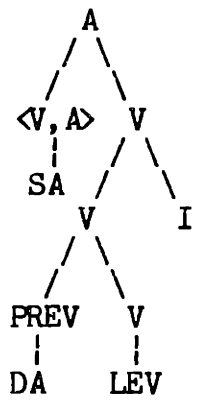
(140)



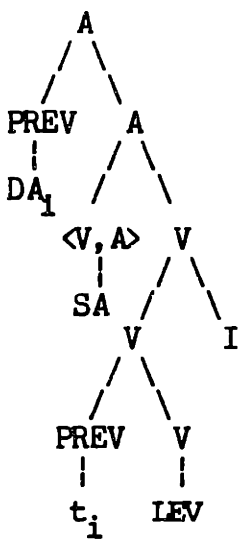
In our terms, we could even postulate that the Move-alpha relation is between D-structure and S-structure and that at D-structurethe preverb and the verb are sisters, and that at S-structure the preverb is adjoined to the highest X^0 category dominating the verb. For dasalevi this would imply the following two structures:

(141)

D-structure:



S-structure:



Note in particular, that this is not a mapping between S-structure and LF.

Be that as it may, it may be thought that in being forced to posit such an operation of Move-alpha in morphological forms we end up with a weaker theory of morphology. Perhaps this is so, but it may be that we can get away with weakening the theory only in selected places. Notice, in fact, that the parallel situation to the Georgian case exists with respect to the verb particle constructions in English: although, say, look and up are associated idiosyncratically in the verb look up, it is nevertheless the case that they are separated in many constructions, including constructions with category-changing morphology:

(142)

a. The book was look$_V$ed$_V$ up.

b. The book remained unlook$_V$ed$_A$ up.

c. The look$_V$ing$_N$ up of the book.

(But contrast this with the -er examples discussed above.) It is presumably no coincidence, too, that preverbs in Georgian, as well as in many other languages, like the particles in English, are related to directional prepositions. So, according to Aronson (p. 42), da- can mean 'down,' whereas another preverb, ga-, means 'out, away.'

It may be then that the Georgian and English cases represent a natural

class of phenomena, a case of universally discontinuous morphology. Why this discontinuity should be characteristic of this particular class of cases, I do not know, but I hypothesize that such cases of discontinuity, when found, will be limited in this way to a fairly well-defined morphological class.

1.6 Conclusions and a Prospectus

I have argued in this chapter that word structure is best viewed as consisting of two levels of representation, namely a syntactic level and a phonological level, and that many of the so-called Morphological Bracketing Paradoxes should be seen as falling out as a consequence of the mapping between these two levels of representation.

Having split morphology in this way, it now remains to determine the nature of the principles which determine well-formedness of word-structure at the various levels. As I mentioned earlier, I am not going to be very concerned with the status of morphology which interacts heavily with syntax in the sense of Marantz (1981, 1984a) or Baker (1985c). I shall be looking more closely at the kinds of principles which determine well-formedness of more derivational/lexical morphology. As noted in the introduction to the thesis, in the next Chapter, I shall be looking fairly closely at deverbal morphology in English and other languages, with a view especially to

determining the effects of the nominalizing affixes on the thematic structure of the base verb. In Chapter 3, I shall be concerned with the so-called Lexical Integrity Hypothesis and its supposed role in deriving the Anaphoric Islandhood of words. In Chapter 4, I turn to a discussion of the stratum ordering model of phonology. Finally, in Chapter 5, I discuss the psycholinguistic evidence which bears on morphological theory, and argue that it is at least as compatible with the approach to morphology taken here as it is with other approaches.

Appendix

Some Further Consequences of the Mapping Principle.

1 A Further Consequence of the Mathematical Properties of * and ^.

Suppose that we have a syntactic bracketing [[A B]C]. Suppose further that A is spelled out phonologically as a prefix--call it **a** and that C is spelled out as an unordered affix--e.g. as a CV template as in Arabic--call it **c**. Also, let **b** be the phonological spellout of B. Then it will follow that PH([[A B]C]) must be **((a^b)*c)**. In particular, because of the non-associativity of *, it will follow directly that this cannot rebracket to **(a^(b*c))**.

To give a hypothetical example, suppose that a language has a prefix T, whose phonological form is t-, which is equivalent to English un- in that it syntactically attaches only to adjectives and forms adjectives from them. Suppose further that there is a deadjectival nominal which phonologically involves providing a CV template of the form CCVC; let us

call this affix X. Finally, suppose that there is an adjective of the form kan which can be affixed with X to form kkan and with T to form tkan. What will be the phonological form of the following bracketing?

(1) $[[T_{\langle A,O\rangle}\ KAN]_A\ X_{\langle A,N\rangle}]_N$

Given the specification that t- is a prefix and that CCVC is (clearly) not ordered with respect to its base, we will derive the following via the Mapping Principle:

(2) ((t^kan)*CCVC)

This will, given the non-associativity of *, become:

(3)

```
t k a n
| | | |
C C V C
```

That is, CCVC affixes to the whole of tkan. What cannot happen is for CCVC to associate only with kan, to yield the following as final output:

(4)

```
t k   a n
| |\  | |
C C C V C
```

(We assume in this hypothetical case that such a representation would not be ruled out on phonotactic grounds.)

Unfortunately, although this prediction is fairly clear, I know of no

data which directly bear on this point. For one thing, in languages such as Semitic which do make use of CV-templates heavily, there is not much relevant prefixal or suffixal morphology to allow one to test out this consequence of the theory. This question therefore awaits further data.

2 On Examples of Stratum I Prefixes outside Stratum II suffixes in English.

Strauss (1982b) has argued (pp. 40-42) that the Stratum I prefix en- attaches outside the Stratum II suffix -en in examples like embolden, enlighten, and so on; the reader is referred to Strauss' discussion for arguments to the effect that this is in fact what is going on. Apart from this possible case, examples of Stratum I prefixes outside Stratum II suffixes do not exactly abound. Part of the reason for this is undoubtedly that there are very few Stratum I prefixes which are sufficiently productive and semantically transparent so as to allow attachment to basically anything that has the right syntactic and semantic properties; remember from the discussion in Chapter I that semantic and syntactic productivity were argued to be crucial in allowing Bracketing Paradoxes to occur.

One other example besides Strauss' does exist, however, at least in slightly earlier forms of English. Those are cases in which the negative

prefix in- has apparently attached syntactically outside the adjectival participial suffix -ed. A large number of such cases are listed in Marchand (1969, pp. 169-70):

(5) illimited, inelaborated, inconfused, inexhausted, inextended, inedited, incircumcised, incivilized, incompared, incomposed, inconcerned, inconnected, incontrolled, incultivated, indigested, indiscussed, indisputed, inexpected, informed (= unformed), insuspected.

(Note that inexperienced is the only such example still in common use.) These presumably have the syntactic bracketing:

(6) [IN [V ED]]

Now, given that in the earlier forms of English -ed was Stratum II, and in- was Stratum I, we predict the following phonological bracketing:

(7) [[in v] ed]

In fact, given that in- attaches only to latinate forms phonologically, it ought to follow that such constructions are only possible with latinate base verbs, which in- can thus be adjacent to and hence bracket with. In fact, this seems to be precisely correct in that all of the verbs in Marchand's list are latinate. Indeed, even in present-day English, the examples given in Marchand seem fairly plausible, if not entirely pristine. However, examples of the following form seem utterly implausible:

(8) *ingrandstanded, *imbuttered, *inkilled, *inwaylaid, *indogeared, *imbeaten, *inwatched, *inwritten

Independently of this, I assume that the reason that Marchand's examples are not in use today is that _-ed_ adjectives simply select for _un-_ rather than _in-_.

3 Another Possible Example of Morpheme Doubling.

In the previous chapter I discussed the doubling of the morpheme _-er_ in English forms such as _picker-upper_ and _giver-upper_. Here I give another example from Berber of what may turn out to be the same phenomenon[1].

In Berber when the subject of a clause is extracted, the verb is made into a participle, formed by the prefix _y-_ and the suffix _-n_:

(9)

a. y- čču wryaz aysum
   3sg ate man meat
   "The man ate meat."

b. w- ay- y-čči- n ___ aysum
   who comp eat(perfective root) meat
   _____ _____/
   participle

   "Who ate the meat?"

So, in the (b) example, the subject of the clause has been extracted,

---

1. Thanks to Mohand Guerssel for the data in this section.

leaving the verb in the participial form. Interestingly, when there is subordination, as with a control verb such as want, and the subject of the (higher) clause is extracted, both the matrix and the subordinate verb may appear in the participial form (note the (b) example):

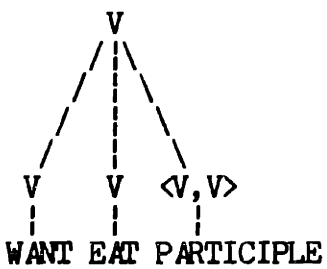
(10)

a. bghi-x ad- čč -x aysum
want-1sg fut eat-1sg meat
"I want to eat meat."

b. w- ay- y-bgha-n _____ ad- y-čč- n _____ aysum
who comp want fut eat(aorist root) meat
___ __/ _____ _____/
participle participle
"Who wants to eat meat?"
(wayybghan later becomes wiggbghan by regular phonological rules)

This is accountable for under the Mapping Principle if we can assume that the two verbs may be reanalyzed to be:

(11)



Some indication, independent of the morpheme doubling facts, that reanalysis must be going on is given by the fact that in cases where the subject is overt, it may occur in the location either of the subject of the

matrix clause, or in the subject of the subordinate clause:

(12)

a. y- bgha wryaz ad- y- ɣč aysum
   3sg want man fut 3sg eat meat
   "The man wants to eat meat."

b. y-bgha ad-y-ɣč wryaz aysum
   "The man wants to eat meat "

The existence of the second example here seems to indicate that the matrix and subordinate verbs have been reanalyzed to form a single verb at some level of representation. We would thus have a structure of the following form for the (b) sentence:

(13) $[_S\ [_V$ want eat] man meat]

Since man in this representation is the subject of the clause headed by want (as well as by eat) we can derive the fact that man is construed as the subject of the higher verb. Presumably, then, (10b) is derived by wh-movement from the subject position of the reanalyzed verb complex. Of course, as (12a) shows, such reanalysis is not required suggesting that we should not always expect to find it in the case of subject wh questions either; indeed, reanalysis (leading to morpheme doubling) is not necessary as the following grammatical variant of (10b) shows:

(14)

w- ay- y-bgha-n _____ ad- y- ɣɛ _____ aysum
who comp want fut 3sg eat meat
___ __/
participle
"Who wants to eat meat?"

Here the wh-item has apparently been extracted from the matrix subject position.

So, this example from Berber might be a further case of the type of morpheme doubling exhibited by -er doubling in English.

Chapter 2

Syntax and Semantics of Productive Deverbal Morphology

2.1 Introduction

In this chapter I investigate the properties of some classes of deverbal morphology, mostly in English, but including discussions of related phenomena in other languages. The framework for semantic interpretation and thematic role assignment which I am going to be assuming is that of Higginbotham (1985b), and I discuss this framework in the next section. I am also going to be assuming some recent work by Chomsky (1984), which I shall also discuss briefly in the next section. After this background I shall be investigating (Section 3.) the properties of agentive *-er* affixation in English, and comparing my approach with a recent proposal by Lieber (1983) for handling these forms. In the subsequent section I shall be investigating the rather complex set of properties associated with derived nominals. Finally, in sections 5 and 6 I shall discuss extensions of my model to other types of morphology, including some facts of

deadjectival morphology of English.

It is important to keep in mind the reason why the investigation in this chapter is necessary for my purposes. I have claimed that there is no separate word-formation component and that the properties of words at various levels of representation should follow from the various principles of grammar which apply at those levels. If we wish to maintain this view we are encumbered with the responsibility to at least initiate the investigation into what the relevant principles of the grammar might be, and furthermore what the properties of the various lexical items--i.e., stems, roots, affixes and so forth--must be in order for the system to work. For the lexical processes which I shall be investigating here, it will turn out that the relevant principles for determining well-formedness will have very little that is lexical about them. Needless to say, this highlights one point which has always, in my opinion, been mysterious in all of the recent work on morphology within generative grammar: apart from the obvious case of Lexical Phonology, as distinct from Post-Lexical Phonology, what precisely does the term "lexical" mean? Given, as has often been noted (c.f. Lieber, 1980; Selkirk, 1982) that there is a "syntax of words", what precisely does the lexicalness of this syntax consist in? I shall suggest here that the answer is "nothing at all."

Note that a massive distinction between lexical compositions and syntactic compositions is certainly conceivable. In particular, certain

operations such as the internalization or externalization of arguments might be argued to be available specifically for forms derived in the lexicon; indeed Williams (1981b) has made precisely such an argument, namely that the derivation of lexical passives in English involves the externalization of the internal theta role. Nevertheless this is not obviously correct, and Levin and Rappaport (1984) in particular have argued that the externalization of the internal argument in the lexical passive follows from independent principles.

2.2 Background: Higginbotham (1985b) and Chomsky (1984).

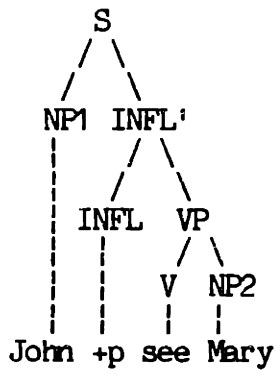
2.2.1 Higginbotham's Semantics.

In this section I give a brief introduction to the theory of semantic interpretation as outlined in recent work by Higginbotham. It is not my purpose to justify any of the assumptions implicit in this system here, as that would lead us too far astray, but rather to give an overview of the main points which are going to be crucial for my discussion of deverbal morphology. The reader is referred to Higginbotham's paper for the relevant arguments.

To begin, let us consider the following phrase marker for the sentence 'John saw Mary', where we will take +p to be an indication on INFL that the

sentence is past tense:

(1)



We will consider the set of nodes N={S, NP, INFL', INFL, VP, V, NP} to be a set of points, these points being related to one another in the obvious way by operators marking dominance and linear precedence (for a discussion of the formalization of these notions see Higginbotham (1985a)). In fact, a phrase marker is precisely just such a set of points like N, and any appropriate subset of N will also be a phrase marker, "appropriate" here referring to whether or not the particular subset chosen is a constituent of the entire phrase marker, as determined by the relation of (exhaustive) dominance. So, in N, {VP, V, NP2} is a phrase-marker, although {VP, V, NP1} is not since VP does not dominate NP1.

The task of semantic theory, according to Higginbotham (1985b), is to determine the values of various points in phrase markers. In particular, we are interested in the behavior of the primitive predicate **v**, which has five places, read as follows:

(2)

$v(\underline{x}, \underline{P}\ \underline{Q}, \underline{C}, \underline{f})$

$\underline{x}$ is a value of the phrase marker $\underline{P}$, considered as a sub-phrase marker of the phrase marker $\underline{Q}$, in context $\underline{C}$, under assignment $\underline{f}$.

In general we will only consider the first two places of v, the others being implicit only.

A simple task of semantic theory to determine that for English, the following statement holds: 'John saw Mary' is true, if and only if John saw Mary. We turn now to a discussion of this example. First of all, we will need to introduce the notion of thematic grid.

Following Stowell (1981), Higginbotham adopts the concept of thematic grid, which is part of the lexical entry of every word and which determines the number of arguments the word can take and also, given that the slots in the grid are marked with thematic information about role, the nature of those arguments. For example, the word see will have the following entry:

(3) 'see', +V -N, <1, 2, E>

This says that see is a verb, and has three places. Although it is not marked in (3), the 1 argument is the designated external argument following Williams (1980), and the 2 argument is the argument assigned by the verb to its internal argument, Mary in this example. Also not marked is the thematic information such as the fact that the 1 argument corresponds to the thematic role AGENT and the 2 argument the role of THEME (following,

originally, Gruber, 1965, and subsequently many others); this information would be present in the full-fledged entry for the verb see. The other argument of see, the one labelled E is the event place. which Higginbotham assumes is present in the representation of all verbs, following an original suggestion of Donald Davidson (1966) that verbs of change or action have such an event position. Thus, the meaning of the sentence 'John saw Mary' will be the assertion that there was an event of seeing perpetrated by John on Mary. The reader is referred to Higginbotham and also Davidson's original paper for justification of this position.

To see the notion of value in action, let us examine the value of the point V in the sentence 'John saw Mary', diagrammed above. The value of this point is obviously completely determined (modulo the context and the value assignment) by the lexical entry of see. In fact, it should be given as follows:

(4) $v(s,V) \Leftrightarrow (\exists x)(\exists y)(\exists e)(s=\langle x,y,e\rangle) \& see(x,y,e)$

What this says is that the value of the node V is that set of ordered triples $\langle x,y,z\rangle$, designated s, where it is the case that for some x and some y, and some event e, e is an event of x's seeing y.

Similarly the value of NP2 would be given as:

(5) $v(y,NP2) \Leftrightarrow y = Mary$

which says simply that y is a value of NP2 if and only if y is Mary.

The value for VP will be given by combining the values of V and of NP2 as follows:

(6) v(⟨x,e⟩, VP) <<=>> (Ey) v(⟨x,y,e⟩, V) & v(y,NP2)

This reads as follows: the value of VP is the set of ordered pairs ⟨x,e⟩ if and only if there is some y such that the set of triples ⟨x,y,e⟩ is a value of V and y is a value of NP2. Given the values for the two statements of the conjunction in (6), given in (4) and (5) above, (6) reduces to:

(7) v(⟨x,e⟩, VP) <<=>> see(x, Mary, e)

so that the value of VP is that set of pairs (x,e) such that e is an event of x's seeing Mary.

Applying this method up the tree we will finally get to a statement of the form:

(8) S is true <<=>> (Ee) see(John, Mary, e)

or, in other words, 'John saw Mary' is true if and only if there is an event of John's seeing Mary.

So much for semantic interpretation. But how do we guarantee that the correct NPs get matched up with the correct thematic roles? This is determined by the syntactic properties of the construction, in conjunction with conditions on thematic role assignment. A central problem for any

theory of grammatical knowledge is the question of how thematic roles (henceforth, theta-roles, following Chomsky (1981)), usually associated with canonical argument-takers such as verbs, are assigned to arguments. In particular, it is generally assumed that there is some form of biuniqueness condition on theta-role assignment to the effect that each theta-role is matched with one argument and each argument is matched with one theta-role. One such biuniqueness condition is the Theta Criterion of Chomsky (1981), given below:

(9) Theta Criterion (Chomsky, 1981)

(i) Each argument bears one and only one theta-role, and
(ii) each theta-role is assigned to one and only one argument[1].

In the work of Higginbotham, this condition takes on a somewhat different, though equivalent, format as we shall see.

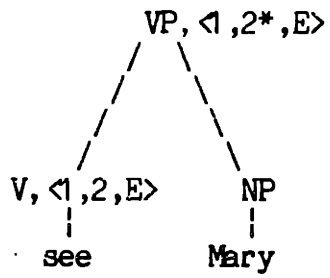
Take again the sentence 'John saw Mary'. The internal theta role of <u>see</u> is assigned to Mary. Following Higginbotham, we will call this particular kind of assignment "theta marking". Theta marking will turn out to be only one way in which theta roles can be <u>discharged</u>. Higginbotham notates theta role discharge in the following way: say that thematic grids are percolated up the tree from the heads, in this case the verb. At each node where a

---

1. This formulation is somewhat inaccurate with respect to current theory (Chomsky, 1984). In particular, the theta criterion is stated in terms of chains where each chain has only one theta position but each theta-position may have a number of theta-roles.

thematic role is discharged we notate that fact by placing an asterisk next to the appropriate role. For instance, the VP 'see Mary' will have the following representation:

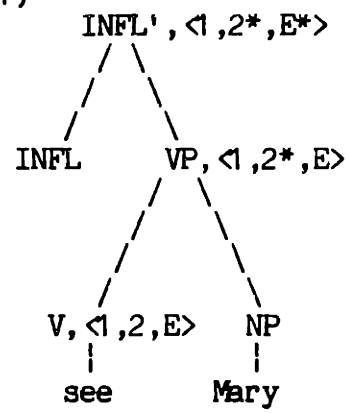
(10)



So, here we have notated the fact that the 2 argument has been discharged to Mary, in this case under theta-marking.

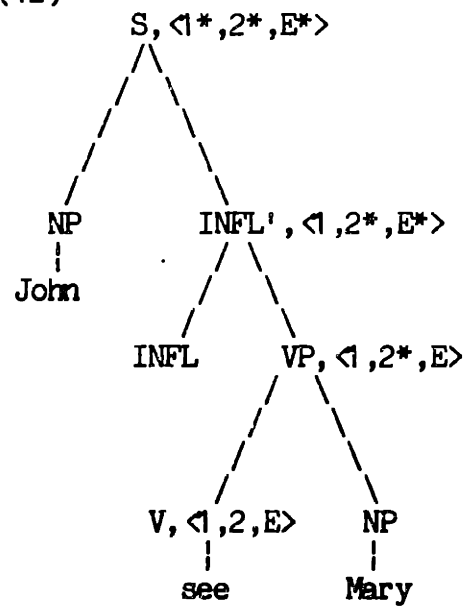
Another way of discharging theta roles is "theta binding". In particular, we say that INFL binds the event position in a theta grid insofar as it restricts the set of events which are permissible referents of the sentence. So, past tense restricts the events to being events in the past in the sense that if 'John saw Mary' is to be true, then it must be the case that there is an event e which is an event of John's seeing Mary, and furthermore, e must have occurred before the time of utterance of the sentence. So INFL discharges E via theta-binding, and the representation of INFL' will be as follows:

(11)



John will finally be theta-marked with the 1 theta role, under predication from the VP. The entire representation will thus appear as follows:

(12)



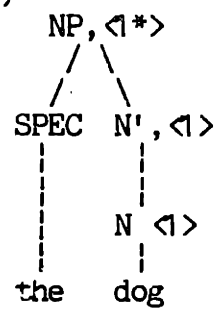
Another case of binding which will become important in the section on -er nominalizations is the binding of a nominal theta position by the SPEC of

NP. Higginbotham suggests that a noun like dog has an open place, or theta position, in it, and hence refers to "each of the various dogs." Its entry will be:

(13) 'dog', -V, +N, <1>

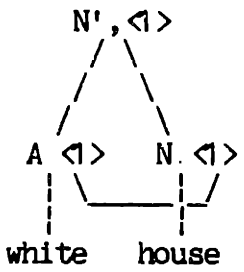
In the NP 'the dog', the as a specifier will bind the 1 role of dog:

(14)



One final method of theta discharge which we will be considering is "theta identification." A particular instance of this is modification, a simple example being the phrase 'white house'. The intuition we want to capture is that 'white house' refers to those entities which are both white and a house. Assuming that both white and house have a theta role, we will say that those theta roles are identified, this identification being notated by a line connecting the two relevant places in the grids:

(15)



The semantics will be given as follows:

(16)

v(x,A) <<=>> white(x)
v(y,N) <<=>> house(y)

Given identification, we can say, x=y, and thus the N' value will be given by:

(17) v(y,N') <<=>> white(y) & house(y)

We have discussed three methods of theta discharge. How does theta discharge relate to the Theta Criterion? Higginbotham essentially translates the Theta Criterion as given in Chomsky (1981) into the following terms:

(18) Theta Criterion (Higginbotham, 1985b)

i) If X discharges a thematic role in Y, then it discharges only one.
ii) Every thematic position is discharged.

I will have cause to refine this formulation in the course of the investigation of the properties of productive deverbal morphology. In particular, (ii) will have to be stated with reference to a reformulation

of the Projection Principle which I shall introduce in Section 2.3 below. For now, however, we have given sufficient introduction to the model of semantics proposed in Higginbotham (1984). I now turn to a brief discussion of some ideas in Chomsky (1984) which will also be important for my analysis.

2.2.2 Chomsky (1984)

In this work, Chomsky introduces some important modifications of various principles discussed in Chomsky (1981). One such modification has to do with the Binding Theory. In the old Binding Theory it was stated that a pronoun had to be free in its minimal governing category whereas an anaphor had to be bound. The minimal governing category (MGC) was defined as the minimal category containing a subject (i.e., [NP, NP] or [NP, S]) and a governor for the pronoun or anaphor. As such, the Binding Theory cannot account for the following facts (see Chomsky, 1984, pp. 240ff):

(19)

a. The children$_i$ like [each other's$_i$ friends]
b. The children$_i$ like [their$_i$ friends]

(19b.) is not a problem since the NP does contain a governor for <u>their</u> and also a subject, namely <u>their</u> itself. The pronoun is free in its MGC. On the other hand (19a.) is problematic precisely because the same reasoning about the structure of the NP applies, yielding the conclusion that <u>each other</u> is free in its MGC, which is an incorrect conclusion since the

sentence would thus be ruled out (on any indexing), yet it is good on precisely the indexing given to it in the example above.

Chomsky thus introduces the notion of BT-compatibility, defined as follows:

(20)
In an expression E with indexing I, I is BT-compatible with (a, b) if:
(A) a is an anaphor and is bound in b under I
(B) a is a pronominal and is free in b under I
(C) a is an r-expression and is free in b under I.

The licensing condition for a category a governed by g in the expression E with indexing I, is introduced as follows:

(21) For some b such that (i) or (ii), I is BT-compatible with (a, b):

(i) a is an r-expression and (A) if a heads its chain or (B) otherwise:

(A) b = E
(B) b is the domain of the head of the chain of a.

(ii) a is an anaphor or pronominal and b is the least CFC containing g for which there is an indexing J BT-compatible with (a, b).

CFC denotes the "Complete Functional Complex", the governing category in which all the "grammatical functions compatible with the head are realized in it" (Chomsky, p. 238), including the subject.

What all of this entails, among other things, is that the two sentences given above in (19) are grammatical. To quote Chomsky (p. 243):

> If a is the subject of the NP b, then it is governed by the head N and b has a subject so it is the Governing Category if there is an indexing I BT-compatible with (a, b). There is such an I if a is a pronominal (namely, with a free) [so (19b) is grammatical: RS] but there is none if a is an anaphor, so in this case the minimal CFC containing b is the Governing Category in which a must be bound [namely, the whole S, in the case of (19a), meaning that that sentence is also grammatical with the appropriate binding, RS].

The other modification which will be important for us is that of Case theory. In Chomsky (1981) it was assumed that the only Case assigners were V, INFL and P. In the newer version Case assignment is split into two categories, Inherent Case assignment and Structural Case assignment. Inherent Case assignment occurs at D-structure under government and theta-marking from P, N or A. Structural Case assignment, on the other hand, is assigned, purely under government, at S-structure from V or INFL containing AGR. The reason theta-marking is not crucial in this case is in order to account for, on the one hand, Exceptional Case Marking, which occurs under government but not theta-marking, and on the other, the fact that INFL assigns Case to the subject even though it does not assign it a theta-role. It turns out also to be necessary to distinguish inherent Case assignment, which occurs at D-structure, from the realization of that Case, which occurs at S-structure. One reason has to do with the now common assumption that Case-assignment is directionally uniform (see Koopman, 1984 Travis, 1984; also Sproat, 1985a). In English it is clear that P and V both assign Case rightwards, and the same would seem to be true for N's, as in (22a,b) below, where we can assume that inherent Case is assigned to

the complement of destruction:

(22)
a. the [destruction [the city]]
b. the [destruction of [the city]]
c. [the city$_i$]'s [destruction e_i]

However, as (22c) shows, genitive Case may also appear on an NP to the left of the head. This suggests that there is reason to separate Case assignment at D-structure, from Case realization at S-structure; it is Case assignment which is directional, and in the mapping from D-structure to S-structure, an NP which has been assigned inherent Case may move to a non-Case position.

(Even though we will assume that this characterization is more or less correct, one cannot help noticing that there are problems with maintaining the view that Case assignment is uniform in English. For one thing, while the analysis worked for (22c) above where it could be convincingly argued that the NP the city was moved from its D-structure position in the complement of the NP to the NP's subject (c.f. Anderson, 1979), it is not at all clear how we would account for the presence of genitive Case on the NP the Romans in the following example:

(23)
a. [the Romans] [destruction [the city]]
b. [the Romans]' [destruction of [the city]]

In this Case, since the NP the Romans is the subject of the larger NP, it makes sense to assume that it started out in the initial position of the

NP, but if this is true, it must be the case that inherent Case assignment may occur leftwards after all. The other sticky point in the analysis is INFL which might appear to Case assign to the left; see Sproat (1985a) for an analysis of VSO languages based on the premise that in VSO languages, INFL Case assigns rightward along with the categories V and P, whereas in SVO languages, INFL is not constrained to Case-assign in any particular direction[2].)

Assuming that Chomsky's distinction between Case assignment and Case realization is a valid one, and is motivated by the data discussed above, it turns out to be necessary to assume that the "association of inherent Case and theta-marking extends to Case-realization as well as Case-assignment" (Chomsky, 1984, p. 271.) He proposes the following condition:

(24) Uniformity Condition, Chomsky (1984)

> If **a** is an inherent Case-marker, then **a** Case-marks NP if and only if **a** theta-marks the chain headed by NP.

"Case-marking" in this case refers to both Case-assignment and

---

2. Still, as Noam Chomsky has pointed out to me, we could use the reasoning in Travis (1984) to the effect that in general not all directionality requirements are satisfiable at once. In particular, for the Romans' destruction of the city, rightward Case assignment, (assuming also Case adjacency (Stowell, 1981), can only occur to one of the NPs. If this is restricted, say, to the internal argument, then the Romans will, in order to also be adjacent to its Case assigning noun, have to appear to the left.

Case-realization. This condition turns out to be necessary, in part, to block the realization of inherent Case on NPs which do not receive theta-roles, for instance, there in the following example:

(25) *there('s) destruction of the city

Since there is an expletive, it cannot be receiving a theta-role from the head destruction. By the Uniformity Condition it cannot be Case marked by destruction either, which means that realization of genitive Case on it is illegitimate. On the other hand, there must get Case since as an NP with phonetic content it is subject to the Case Filter. Therefore, the structure is out altogether.

We will return to the matters introduced here in subsequent sections.

2.3 The Formation of Agent Nominals.

2.3.1 Introduction.

In this section we investigate the syntactic and semantic properties of a particularly productive morphological operation, that of Agent Nominal formation. Examples of the kinds of constructions which we will be investigating are given below:

(26)
a. John is a cookie-maker.
b. John is a maker of cookies.
c. *They$_i$ saw three admirers of{each other$_i$ / themselves$_i$}
d. They$_i$ saw each others'$_i$ admirers.
e. Mark is a driver.
f. *Mark is a driver in a car.

The citation of some of these examples at this point may seem somewhat mysterious. The mystery should clear up in subsequent discussion.

The interesting thing about Agent Nominalization cross-linguistically is that it is often the only productive process of Nominalization in many languages. Also, Noun-AgentNoun compounds are often the only compounds in the language. The latter is more or less the situation in Chichewa, where Agent nominalization, by means of a prefix is productive as are compounds formed with these:

(27)
a. sodza 'fish (V)'; msodzi 'fisherman'
b. chi- lima- mpunga 'rice cultivator'
CLASS cultivate rice
c. m khala pa mwala 'rock sitter'
CLASS sit on rock

The gloss of the prefixes chi- and m- as AGENT would be inaccurate since these are really just two of the many noun class prefixes common in this language, as in other Bantu languages, hence the designation in the gloss. Note that (27c) really is a compound and has the structure:

(28)

m	[khala	pa mwala]
CLASS	sit	on rock

I shall argue for this structure for Chichewa agent nominals later in this section (2.3.7). For now, I turn to a discussion of the semantics and syntax of the -er affix in English.

2.3.2 -Er affixation

Let us start off by asking the following question: Given that Agent Nominal formation is semantically transparent as it clearly is, what is the semantic contribution of the -er affix to the word which it forms? That is, given the meaning of the verbal stem what meaning do we have to assume is given for -er such that the meaning of the V-er noun is predictable given these two parts plus the rules of semantic composition?

Agent nouns refer in general to persons or things which typically engage in some activity or other. That is, a driver is one who typically engages in an activity of driving, and a catcher is one who typically engages in an activity of catching. What we want to say (following a suggestion to me by Alec Marantz) is that -er refers to an Actor in an event. Actor is a term suggested originally by Gruber to mean what is generally referred to nowadays as Agent. Let us extend the meaning of Actor here to mean any "agenty" sort of role such as Agent or Instrumental. This Actorhood of -er can be seen most clearly, in fact, in certain cases where -er attaches

irregularly to nouns. Some examples are given below:

(29) footballer, carpet-bagger, messanger....

A footballer, of course, is someone who plays football. There is however no verb *to football meaning 'to play football,' and it is clear that the structure of this word is something like:

(30) $[\ \text{FOOTBALL}_N\ \text{ER}]_N$

Nevertheless the Actorial meaning associated with -er is clear since footballer refers to an Actor in an event of football-playing. A similar story holds of the other examples: carpet-bagger is not derived from a verb but rather from the noun-noun compound carpet-bag, the referent of this term being, of course, a particular group of individuals of somewhat questionable moral status who invaded the South from the North after the American Civil War for the purpose of making good at the expense of the war-ravished South, and who could often be identified by the carpet-bags which they carried. In this case, again, the -er imparts an Actorial meaning, the specific meaning being something like 'one who carries a carpet-bag.'

I suggest, then, that productive -er has the following lexical entry:

(31) $\text{ER}' = \langle \text{ER}_{\langle V,N\rangle}\langle 1\rangle,\ \text{er}\rangle$
suf(er)
Actor(1)

I have notated the theta-role of -er on the syntactic half (i.e., the left

half) of the entry for the affix.

In the case of regular affixation of _-er_ to verbs we want to say that the 1 role of _-er_ is linked up with the external theta role of the verb to which it attaches, in case that theta role specifies an Actor of some kind (i.e., Agent, or possibly Instrument). This restriction on the thematic role associated with the external theta role of the verb seems necessary since there are verbs which have non-Actorial external theta roles which do not have corresponding _-er_ nominals. One such case is the verb _surround_

(32)
a. The trees surround the house.
b. *The trees are surrounders of houses.

In this case the external theta role of _surround_ is a Theme, and is hence inappropriate for the linking to the role of _-er_.

But precisely how does the 1 role of the affix link to the 1 role of the verb? To answer this, let us consider the notion "External Argument," a notion which was introduced in Williams (1980, 1981). The External Argument is, by definition, the last argument (i.e., outermost) to be satisfied in any syntactic structure. This is clear in an example like the following:

(33) Josiah Q. Smedley put the schnauzer into the box

Put satisfies its two obligatory internal arguments inside the VP, and only then does the entire VP assign the external theta role to the subject NP _Josiah Q. Smedley_. We shall see below a very similar behavior in derived

nominals, although in that case there is no maximal projection present to act as a predicate for the subject.

Of course, it is also possible for the External Theta Role to be dispensed with entirely, as in the so-called middle constructions as discussed by Keyser and Roeper (1984):

(34)
a. Bureaucrats bribe easily.
b. This meat chews well.
c. That car drives comfortably.

Under normal circumstances, given that the external thematic roles are assigned internal to the VP, and given a theory of Predication, such as that of Rothstein (1983), which claims that Predicates must have subjects, it will turn out that the VP must assign some thematic role and since there is nothing else around at this stage the External Theta Role will be the one assigned. However, Keyser and Roeper's data show that in certain cases the external theta role may be dispensed with.

Let us therefore take the following definition as a guiding assumption:

(35) Definition of External Theta Role.

> The External Theta Role, if it is assigned at all, is always the last (i.e. outermost) role to get assigned in any syntactic structure.

The determination of whether, in some specific case, The External Theta Role must be assigned is determined by other principles of the grammar such as the principle of predication mentioned above. In terms of

Higginbotham's theory let us reformulate this statement as follows:

(36) Definition of External Theta Role (Reformulated.)

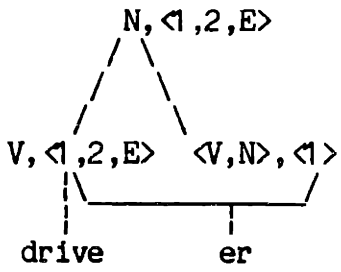
> The External Theta Role, if it is discharged at all, is always the last (i.e. outermost) role to get discharged in any syntactic structure.

Given this, it is clear that the External Theta Role of the verb cannot be discharged to the affix, unless perchance all other theta roles of the verb are discharged at the point at which -er attaches. But this will never be the case since even in monadic predicates, there will still be the Event theta role, which would be undischarged at the relevant level (What does happen to the Event will be discussed below.)

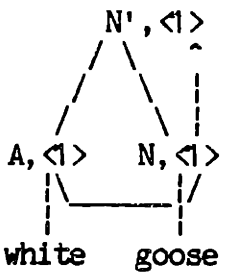
Let me suggest, then, that the theta role of the affix is discharged under identification to the External Theta Role of the verb, assuming, of course, that the verb's External Theta Role is of the appropriate thematic type. This means that although the affix is the syntactic head, the verb will act as the semantic "head" of the construction, and will percolate its theta grid. Note that this is similar to what happens when an adjective discharges its theta role under identification with the noun, with the noun percolating its theta grid, with the difference being that the noun is also the syntactic head:

(37)

a.



b.

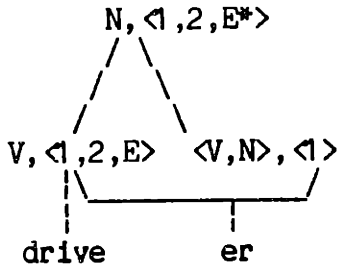


In effect, then, -er acts as a predicate modifying (in purely formal terms) the verb to which it attaches.

We will investigate what happens to the external and internal theta roles of the verb shortly. What happens, however, to the Event? Nouns do not canonically allow event positions. That is to say, there is no sense in which dog has a free event position refering to a particular event of being a dog. Nouns may well refer to events, as we shall see in the discussion of derived nominals in due course, but they do not carry a distinguished event position. Something therefore has to be done with this position in the percolation across the N category boundary. Let us suggest that in this case the event position is existentially bound. We will notate this in the diagrams by starring the position of the Event in the

noun as usual, though we may equally well say that the Event position is simply absent at the N node:

(38)



What will the semantics of these forms be then? For V and <V,N> we can say the following:

(39) $\mathbf{v}(\underline{s},\underline{V})$ <<=>> $(\underline{E}\underline{x})$ $(\underline{E}\underline{y})$ $(\underline{E}\underline{e})$ $(\underline{s}=\langle\underline{x},\underline{y}\ \underline{e}\rangle)$ & drive$(\underline{x},\underline{y},\underline{e})$

(40) $\mathbf{v}(\underline{z},\ \underline{\langle V,N\rangle})$ <<=>> Actor$(\underline{z})$

Given this, and given that $\underline{z}$ and $\underline{x}$ are identified, we get the following result for $\underline{N}$:

(41) $\mathbf{v}(\langle\underline{z},\underline{y}\rangle,\underline{N})$ <<=>> Actor$(\underline{z})$ & $(\underline{E}\underline{e})$ [drive$(\underline{z},\underline{y},\underline{e})$]

That is, <u>driver</u> refers to any Actor $\underline{z}$ for which there is an event $\underline{e}$ which is an event of $\underline{z}$'s driving some $\underline{y}$ or other. This is the right interpretation.

I should point out at this juncture that languages other than English seem to emphasize slightly different aspects of Actorhood when forming

Agent nominals. For instance, in Kalkatungu (Blake, 1969), a language spoken in Queensland, the morpheme which translates English -er with the same morpheme as it uses for can, to be able.' Thus -ŋcit can be used both as a verbal affix, which we may assume is attached syntactically, or as a deverbal nominalizing affix, as shown by the following examples (Blake, p. 76):

(42)

a. cutu caa kuntu palku tuŋaŋcit
coolamon that not slow run-can
"a coolamon that can run not slow" = "car"

b. caa miltipia unpicin naŋintiicit
that eyes-on take-ing see-make-can
"that which, when taken upon the eyes, can make one see" = "spectacles"

c. kaniŋaŋcit
knot-make-er
"policeman"

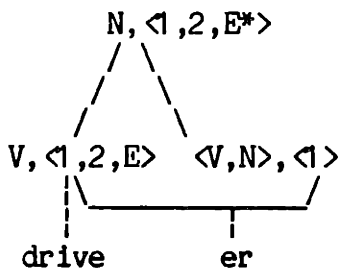
d. wakini-wakini kuu anpaicit
turn turn water fetch-er
"windmill"

Evidently, in Kalkatungu, it is the general propensity towards doing a particular action which is being emphasized by use of the morpheme -ncit. In Chichewa, as noted above, and as I shall discuss more fully below in this section, the affixes corresponding to English -er are Noun Class affixes, affixes which are used in a much wider range of Chichewa morphological structures than just in forming deverbal nominals. It might

seem that it would make little sense to assume that the entry for these affixes includes a specification that there is a thematic role associated with the affix, and that this role must be hooked up with an Actor. In fact, I shall suggest below that in Chichewa, the "agent" affixes are in fact nothing more than nominalizers. This will be shown to derive an interesting consequence for that language.

Getting back to English, we have seen that the regular rules of thematic linking and semantic interpretation allow us to derive a form such as driver which has the correct meaning. The final form of this word is repeated below:

(43)



The 1 role of verb has become the derived noun's own thematic role; we may assume that this is the case since it is the one which is linked, after all, to the syntactic head of the construction. The 2 role is still free, and presumably can be assigned. This, of course, is correct, as the following forms show:

(44)
a. John is a driver of large trucks.
b. Hortense is an admirer of pink aardvarks.

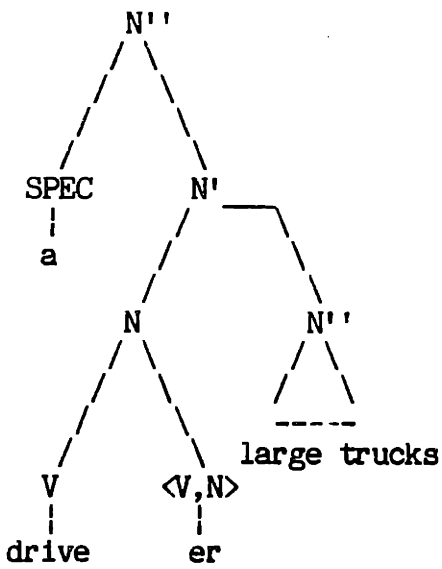
c. Edgar is a lover of Hunan cuisine.
d. Peter is a painter of lurid landscapes

So, in (44a.) the NP large trucks is getting the theta role--the 2--which would normally be assigned by the verb to its internal argument. Thus, (44a.) has roughly the same meaning as:

(45) John is a person who drives large trucks.

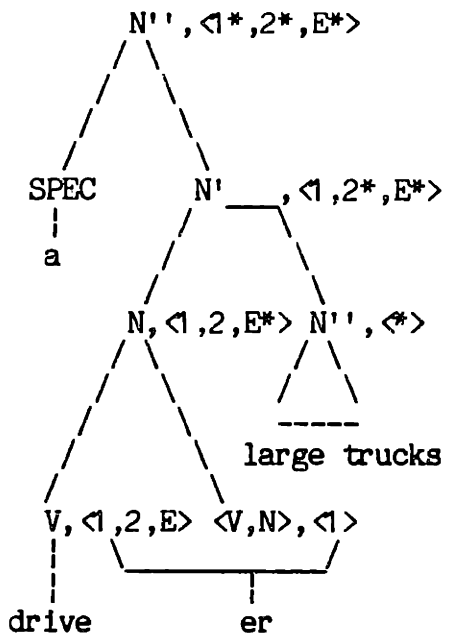
We can derive this fact, as follows. We may assume that the relevant syntactic structure of (44a.) is:

(46)



Filling in the thematic roles, we have:

(47)



This will have the following interpretation:

(48) v(x, NP) <<=>> x=a z[Ee, drive(z, large trucks, e)]

Thus, a driver of large trucks is some z such that there is an event e, such that e is an event of driving large-trucks by z.

Of course, we fully expect that a verb which requires an internal argument will also yield an -er nominal which requires such an argument. This is correct:

(49)
a. John put the book on the table.
b. John is a putter of books onto tables.

c. *John put the book.
d. *John is a putter of books.

e. *John put on the table.
f. *John is a putter onto tables.

(50)
a. Bill makes cookies.
b. Bill is a maker of cookies.

a. *Bill makes.
b. *Bill is a maker.

Furthermore, if a verb has certain of its internal arguments optional, or if a verb alternates between being transitive and intransitive, we expect these possibilities to show up in the _-er_ nominals also. A fairly clear example of this is given below:

(51)
a. Fred drinks tea.
b. Fred is a drinker of tea.

c. Fred drinks.
d. Fred is a drinker.

In (51a,b) _drink_ assigns an internal theta-role, in this case to the object _tea_. In (51c), on the other hand, _drink_ is intransitive, and has the very specific meaning 'imbibe alcoholic beverages. As (51d) shows, the intransitivity carries over to the _-er_ nominal. Indeed, the particular meaning 'imbibe alcoholic beverages' also carries over. That is, to say that Fred is a drinker is not to be making the claim that Fred, like all of the rest of us, drinks liquids, but rather to be making the claim that Fred, like only some of the rest of us, is a regular imbiber of various ethanol products. We may assume that the following way of representing the two readings of _drink_ is appropriate:

(52) Two (syntactic/semantic) entries

i) DRINK, V, <1,2,E> '1 imbibes 2...'
ii) DRINK, V, <1,E> '1 imbibes alcoholic beverages...'

In (52) we are essentially claiming that there are two verbs drink. Under this scheme the information as to which verb is being nominalized will be carried over to the -er nominal by virtue of the percolation of the grid from the verb to the nominal, as the reader will be able to easily verify at this point.

There is another consequence of this theory. Randall (1984) has discussed what she calls the Inheritance of theta roles. She notes that when certain morphological operations appear to block the assignment of various thematic roles of the stems, usually verbs, to which they apply, they also seem to block assignment of thematic roles lower on the Thematic Hierarchy, which she takes to be:

(53) Theme>Agent>Instrument, Source, Goal.

A particularly clear case of this involves -er affixation. This affix blocks the assignment of the external theta role, invariably the agent (Randall, p. 74):

(54) the flyer of the kite *by experts

This makes sense of course under our approach since the nominal which is produced is the agent, and the External Theta Role of the verb can only be assigned to the referent of the NP headed by the -er nominal, under binding

from SPEC. What is interesting for Randall is that other thematic roles such as Instrument or Direction seem to be blocked also[3]:

(55)
a. *the flyer of the plane to France
b. *the cutter of stamps off of envelopes
c. *the destroyer of the city with a machete

(I should note in passing that I disagree with many of Randall's grammaticality judgements, especially with regard to obligatorily subcategorized PPs. For me a paster of stamps onto envelopes is fine, despite Randall's claim to the contrary, and this makes sense if we assume, as I believe we must, that paste, at least on one reading, requires two internal theta roles, one of which is expressed by a PP[4]. In any event, I agree with Randall's judgments in the cases of the non-subcategorized PPs.) To account for these data, she states the following principle:

---

3. As pointed out to me by Alessandra Giorgi, sentential complements are consistently bad here: *A believer that the world is flat. I have no explanation for this fact.

4. Tom Roeper has pointed out the following examples to me:

(i) *a letter of dogs into houses
(ii) *a keeper of people in jail

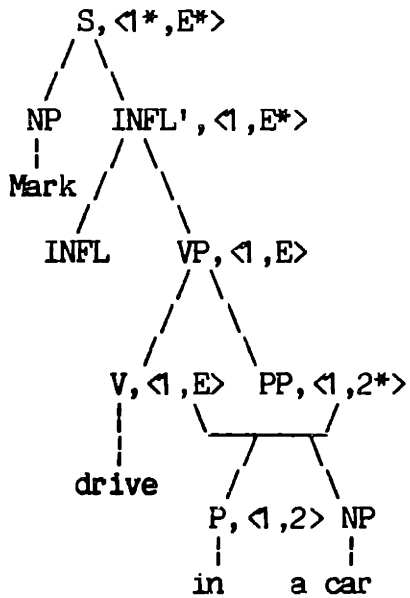
where surely the direct object and the prepositional object are both subcategorized for. It may be the case, then, that even many subcategorized PPs are obligatorily linked to the event position (see my account for Randall's facts below.) If this were true of into houses in (i), this would account for why the example is apparently bad.

(56) Thematic Inheritance Principle, (Randall, 1984, p. 77)
A category-changing operation which blocks the assignment of a θ-role blocks the assignment of all θ-roles lower on the θ-Hierarchy.

So, in the case of *-er* affixation, we block the assignment of the Agent, and we therefore also block the assignment of the Instrument, Direction and so forth.

But this fact is derivable under the theory presented here without resorting to a Thematic Inheritance Principle applying specifically to lexical items. It is assumed crucially (Higginbotham, 1985b, and p.c.) that many verb-phrasal PPs act as adverbial modifiers of the Event. For instance:

(57)

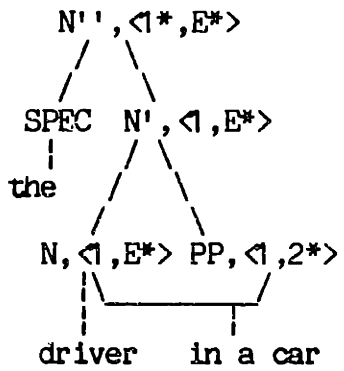


So in *Mark drives in a car* the PP *in a car* has an open position which it

discharges to the event position of the verb under theta identification, thus modifying it, and giving the reading that there is some event e, where e is a driving by Mark, and e is in a car (see also Davidson, 1966.)

Given this, the Thematic Inheritance Principle effects can be derived straightforwardly for -er nominals since in those cases, as for most nouns, there is no event. The structure would be:

(58)



The 1 position of the PP needs to link to something but it cannot, in particular, link to the E position since that is already discharged. Therefore the NP cannot refer to someone who is in the habit of driving in a car. Of course, the 1 position can link to something, and as diagrammed in (58), it links to the 1 position of the noun, yielding the reference for the whole NP of a person who drives and who is in a car, which is, of course, what the NP means.

We have seen then how part of Randall's Thematic Inheritance Principle falls out naturally from the notation adopted here. We will discuss more

of this principle in subsequent sections and it will turn out (see Section 2.7) that it is completely derivable and hence is not a separate principle of the grammar.

We now turn to an apparent problem for the theory of _-er_ affixation developed so far.

2.3.3 Projection, Projection, and Theta Indexing.

Despite the obvious productivity of _-er_ affixation in English, and the concomitant semantic transparency of such forms there seems to be a problem. It will undoubtedly have been noticed that there are _-er_ forms which do not seem to be semantically transparent in the same way as, say, _driver_ is. Some examples are:

(59) fryer = 'chicken suitable for frying', propeller = 'propeller of a ship', computer = 'computing machine', sinker = 'weight used in fishing to sink the line'...

Fabb (1984) lists a number of these. He says (p. 206):

> _-Er_ attaches to verbs and produces nouns; the derived noun has a relation to the verb characteristic of the relation between the external argument and the verb. For example, a 'warbler' is the agent - the external argument of 'warble'. In the following examples we see that the V-er nominal denotes the external argument of the verb, whether that external argument is agent, source, theme or benefactor:

(5.16) external argument is AGENT
gambl-er, strik-er, warbl-er, march-er, race-er
brew-er, lectur-er, plaster-er, teach-er, retriev-er
(5.17) external argument is SOURCE
gush-er, trail-er, creep-er
silenc-er, fertiliz-er, thrill er, cook-er, scrap-er
(5.18) external argument is THEME (ergative verbs)
twist-er, break-er, crack-er, vibrat-or
(5.19) external argument is BENEFACTOR
hear-er, learn-er discover er, inherit-or

The special meaning of -er nouns is derived from the -er suffix, in some way. We can not, however, simply say that the suffix is 'source' or 'agent', and percolate this property to the word, because whether the word is a source, agent, or theme of an action depends on the theta-grid of the underlying verb. The suffix is 'agent' if the external argument of the verb is 'agent', and 'source' if the external argument of the verb is 'source', and 'theme' if the external argument of the verb is 'theme'. . .

What Fabb fails to notice, however, is that, in the vast majority of cases, unless the noun does refer to an "agenty" sort of thing--such as the Agent or Instrument, etc.--it does not inherit the verb's theta grid in the way described in the last section (despite Fabb's claim, p. 209, that in general " Er does not appear to affect the internal arguments in the theta-grid to which it attaches.") Take for instance the word fertilizer, which Fabb claims expresses the Source of the verb fertilize. I do not think, however, that one can refer to 'a fertilizer of the field' unless one is specifically referring to the person who is responsible for the fertilizing. For another example, take discoverer, which Fabb claims is a Benefactor. I do not know in what sense a discoverer is a Benefactor in any case, but it seems fairly clear that if I say that Leif Eiriksson was the discoverer of Vinland, that I am claiming that he actively set out on a

voyage and discovered it and that therefore he is an Agent in Vinland's discovery.

What is even more striking about many of these words is that they are specific in reference. Whereas driver can refer to anyone or anything that drives, or drinker can refer to anyone or anything that drinks, a gusher generally refers to a gushing oil find and twister generally refers only to a tornado. Cooker refers to an instrument of cooking, though not just any instrument; it has to be a gas or electric range and it could not be an open fire even though an open fire can certainly be used to cook things.

In fact, the cases where -er picks up some argument of the verb other than Agent, or Instrument, seem to behave just like some obviously exceptional cases where the -er does not even pick up the external theta role. Fabb gives the following examples (p. 208):
(60) respirator, howler, roaster, romper, confessor, merger

Thus, respirator is not something that respirates, but rather something which aids in respiration. A roaster is a chicken or some other meat suitable for roasting (see also the example fryer, given above.) Again the reading is quite specific. As pointed out to me by Tom Roeper, if I am lying out in the sun on the beach and I get sufficiently burnt on both sides so as to be able to say that I have roasted in the sun, it would nonetheless not be possible to refer to me as a 'roaster.' Similarly, if I am delivering mouth to mouth resuscitation to somebody, and am thus

obviously aiding them in their respiration, I would not be referred to as a 'respirator.' A 'respirator' is a kind of machine used to aid in respiration.

Of course, not all Instrumental products of *-er* affixation are semantically or thematically productive in the sense discussed in the last subsection. So, *propeller* can be used in the following way:

(61) This is the propeller of the ship

However, in the usual reading of this sentence it is clear that the NP *the ship* is not acting as a Theme, in the same way as, say, *the car* is in the NP *the driver of the car*. Rather, *the ship* is the possessor of the propeller, and there is no reason to suppose therefore that *the ship* is being assigned the Theme role originating in the Theta grid of *propel*. To see this, we only need to note that the propellers of the QE2 will still be the propellers of the QE2 even if the engines fail and the passengers are all told to break out the paddles. In this cases the passengers will be the *propellers of the QE2* in the productive sense.

What I suggest for all of the cases discussed in this subsection is that the words have their syntactic representations listed, just as the word *elephantine* has its syntactic representation listed; they are "lexical idioms" in the sense discussed in the last chapter. And bear in mind that also in the sense of the last chapter, there is no implication in this as to their phonological behavior; the forms discussed in this subsection are,

I believe, all phonologically regular. So, a word like propeller would merely be listed as, say

(62) $[\text{PROPEL}_V\ \text{ER}_{\langle V,N\rangle}]_N$

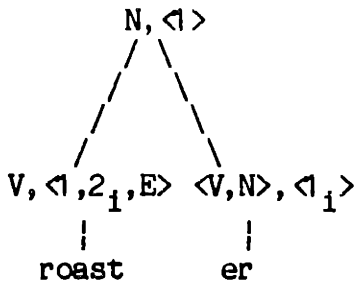
with the listed referent being the fan-like thingie on the back of a ship. Whether or not anything of the Actorial nature of -er is inherited depends upon how semantically transparent the construction is. In the case of propeller, the constuction is reasonably transparent, whereas in the case of roaster it is fairly untransparent. Also, which of the theta roles of the verb the -er affix picks up will depend again on how transparent the construction is; in the case of propeller the theta role is the Instrument, which may at least sometimes act as an external role (c.f., 'This device propels the ship.'), whereas in the case of roaster the theta role picked up by -er is the Theme or Patient role of the verb roast--an internal role.

Even though none of these forms are productive, we want some way of indicating which theta role the -er affix is associated with. To mark this, I introduce theta-indexing, which will freely coindex a position in an (affix's) theta grid with some position in the theta grid of the stem. So, roaster would be represented as follows[5]:

---

5. Jim Higginbotham has suggested to me that theta indexing is really just theta-identification. The representation of roaster would therefore be:

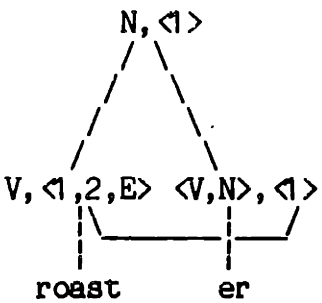
(64)



We may assume that the coindexation of the role of *-er* with that of the Theme of *roast* overrides the specification that *-er* refer to an Actor.[6] Note that this theta-indexing has nothing in common with the similar formalism used in Stowell (1980), or Fabb (1984) and Roeper (1984). In those theories, indexing with a position in a theta-grid was a notation for what we have termed, following Higginbotham, thematic discharge. Here,

---

(63)



Although this is an interesting suggestion, there are details in the implementation of this idea which remain to be worked out.

6. One may wonder why *-er* cannot pick up the Event position, yet so far as I know, there are no nouns of this form. I suppose that there must be some further restriction on *-er* which requires that it refer to an entity of some kind, rather than to an event.

theta indexing is only used to notate the (irregular) association of positions in grids. Note that I have assumed that theta-indexing is only a property of affixes. Thus, obligatorily subcategorizing entities such as nominalizing affixes may elect not to discharge a theta-role of the form to which they attach. This device will be used in the analysis of Chichewa synthetic compouns (see 2.3.7.), where. it is argued, the relevant nominalizing affixes, unlike <u>-er</u> in English, are merely general noun-forming affixes, and hence are not marked specially to identify with a verbal theta-role.

An important property of the diagram given in (64) above is that the thematic grid of the verb does not percolate, unlike the regular cases discussed in the last subsection. Why is this? Obviously we have assumed that there is nothing in principle blocking the percolation of thematic grids across categories. What I will suggest is that thematic grids are required to percolate from stems to dominating nodes of a different category only when the grids are associated with the (grids of) the affixes by some regular method of discharge such as theta marking, theta binding or theta identification. I state this principle as follows:

(65) <u>Cross-Categorial Theta-Grid Percolation Convention</u>

In a structure $[A\ \langle A,B\rangle]_B$ where A is a lexical item of category A <A,B> is an affix forming words of category B from lexical items of category A and the whole structure is thus of category B, the theta-grid of A is only required to percolate to B where A and <A,B> are associated by some regular process of thematic discharge.

However, this brings up the question of how the thematic roles of a lexical item, such as roast in roaster are satisfied. Evidently they are not since none of them are discharged (unless we wanted to say that they were discharged by some existential binding of each of the positions); they are, in effect, "locked up" in the morphology. Yet why is this allowed? Why, given the Theta Criterion are we allowed, as we apparently must be, to ignore the thematic requirements of lexical items in this way? The answer lies, I believe, in the correct interpretation of the Projection Principle, the principle which requires that a position be present at every level of syntactic representation for a thematic role to be discharged to. Its formulation in Chomsky (1981) is as follows:

(66) Projection Principle (Chomsky. 1981.)

> Representations at each syntactic level (i.e., LF and D- and S-structure) are projected from the lexicon, in that they observe the subcategorization properties of lexical items.

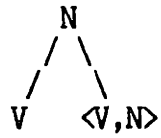
What I would like to suggest is that this notion of projection is linked crucially to another notion of projection, the notion of projection familiar from X-bar theory. Specifically, I would like to suggest that the following principle holds:

(67) Constraint on the Projection Principle.

> The Projection Principle applies to **a** in a structure $[_{\mathbf{g}}...\mathbf{a}....]$ where **g** directly dominates **a**, and **g** is of the same category as a.

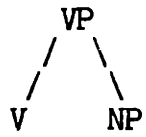
So, in a structure like

(68)



the Projection Principle will not apply to V--since N is [+N,-V] and V is [-N,+V]--in the sense that its thematic roles will not be discharged. It will apply to the affix <V N> since, following Lieber (1980), it is of the category [+N,-V] in that it determines a form of that category. On the other hand, in

(69)



since VP is of category [+V, -N] and dominates (in fact, is projected from) V, the Projection Principle will apply. This conception of the Projection Principle is similar in some respects to Borer's (1984) idea that the Projection Principle does not apply in the lexicon. In my terms, the Projection Principle does not apply to lexical items embedded in certain kinds of structures. Of course, the thematic grids of V in the structure (68) may percolate to the dominating N node and thus be inherited by the noun--i.e. if required to by the Cross-Categorial Theta-Grid Percolation Convention. But this is not to say that the Projection Principle is

satisfied for V, but rather that the N is a form derived from V which has inherited V's thematic grid. Of course, the Projection Principle will apply to N if it heads an NP, requiring that the thematic roles be discharged; that this is the right result has already been shown in Section 2.3.2.

This constraint on the Projection Principle will be crucial in the discussion of synthetic compounding in Section 2.3.5. below. Before we get to a discussion of the properties of synthetic compounds, however, I would like to digress momentarily to discuss some of the interesting binding properties of _-er_ NPs.

2.3.4 On Binding in _-Er_ NPs.

An interesting fact about _-er_ nominals is that although the external theta role of the verb is identified with the derived noun's own theta role, it still acts as a subject for the purposes of the Binding Theory. So, while both of (70a,b) are good, there is a marked contrast for most speakers between (71a,b):

(70)
a. The men$_i$ saw [their$_i$ admirers]
b. The men$_i$ saw [each other's$_i$ admirers]

(71)
a. The men$_i$ saw [admirers of them$_i$]
b. *The men$_i$ saw [admirers of each other$_i$]

This is comparable to the situation with picture nominals with overt subjects (i.e., possessor NPs), though it contrasts markedly with picture nominals with no overt subject:

(72)
a. The men$_i$ saw [their$_i$ pictures]
b. The men$_i$ saw [each other's$_i$ pictures]

(73)
a. The men$_i$ saw [Bill's pictures of them$_i$]
b. *The men$_i$ saw [Bill's pictures of each other$_i$]

(74)
a. The men$_i$ saw [pictures of them$_i$]
b. The men$_i$ saw [pictures of each other$_i$]

We will not be concerned here with the cases where the pronoun or the reciprocal are in the subject position of the NP since these, as we have already noted are taken care of by the new Binding Theory. What is interesting is the contrast between (71b) and (74b). It seems as if there is something within the NP headed by admirer which acts as a subject and hence forms a Complete Functional Complex (Chomsky, 1984; and see section 2.2.2) for the anaphor each other. Each other will then have to be bound within the NP, but the only NP around to bind it would be the entire NP itself, which yields an i-within-i violation (Chomsky, 1981):

(75)
[the [admirers of each other$_i$]]$_i$

What could the subject be? Could it be the head noun admirer itself, a logical guess since that noun does after all pick up the external theta role of the verb admire. The problem comes in examples like (70b) where the reciprocal is in the possessive position of the NP. In this case there will be a governor for the reciprocal and a subject non-identical to it, all of which would imply that the NP should still act as the relevant binding domain for the reciprocal[7].

Say then that the SPEC position, or its contents, acts as the subject of the NP. This is not at all implausible either, insofar as this is the position normally referred to as the subject of NP (see Anderson, 1979, among others.) On the other hand, there is a difference, as we have seen, between the admirer cases and the picture cases insofar as whatever is in SPEC in the former must count as a subject, whereas only overt NPs, or perhaps a "hidden pronominal" (Chomsky, 1984), count as subjects in picture NPs; the point is that picture NPs need not have a subject whereas admirer NPs apparently do. This may not seem so strange, however, when we consider the fact that the SPEC in admirer binds a position in a theta grid which is normal assigned to a subject, whereas the reciprocal is receiving the internal theta role from that same grid. The SPEC is therefore, in some

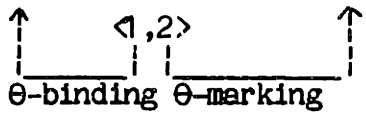
---

7. Though we could assume, as suggested to me by Noam Chomsky, that the whole sentence is still the binding domain in (70b) since the matrix verb see also governs this position.

sense, thematically the subject. The structure of (71b) would thus be, schematically:

(76)

*[The men$_i$ [saw [SPEC [admirers of each other$_i$]]]]



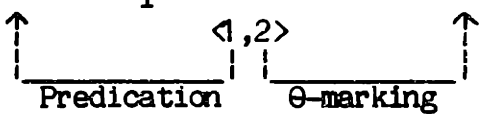
Of course, when an NP headed by an Agent nominal such as admirer is predicated of another NP, the external theta role of the verb will end up being assigned under predication to the other NP. In this case, this external NP will act as the subject, and should therefore be allowed to bind into the Agent NP. This is correct:

(77) I consider John and Mary$_i$ admirers of each other$_i$

This presumably has the structure:

(78)

I [consider [John and Mary]$_i$ [admirers of [each other]$_i$]]



This in fact is in line with the observation (due to Williams and pointed out to me by Jim Higginbotham) that unsaturated categories of which predicates are a particular instance, are not domains for binding.

2.3.5 On Synthetic Compounding.

We now come to one of the central parts of this chapter, the treatment of synthetic compounding with _-er_ nominals. Some of the analysis here will follow that of Lieber (1983) and I shall point out the places where these two approaches agree. I shall postpone a full comparison of the two approaches till a later section.

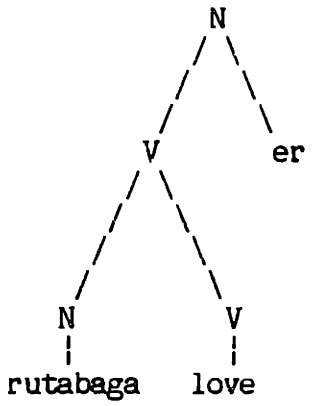
2.3.5.1 The X-bar Status of Synthetic Compounds.

Consider synthetic compounds of the following form:

(79)
a. John is a _city-admirer_
b. Mary avoids _rutabaga-lovers_
c. Charlie is a _worm-hunter_
d. Fred is an _American History teacher_

I shall assume, following Lieber, that the syntactic structure of one of these is more or less as in (80). I shall justify this assumption fully below.

(80)

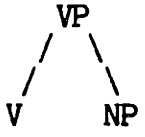


Actually the left-hand member in English must be somewhat more complicated than a simple noun as evidenced by forms such as American History teacher. One thing which is clear, however, is that the left-hand member must not be a full NP, a property which is shared by root compounds. Fabb (1984) gives evidence for this which is quite striking: even proper names which take the definite article as part of their normal structure such as The Bronx or The Hague cannot have these specifiers in compounds. Thus, one can be a Bronx hater, but not a *The Bronx hater.

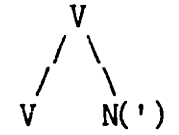
Why should this be so? To ask the question more precisely, why is it that we find structures such as (81a,b) but not (81c) (linear order irrelevant)?

(81)

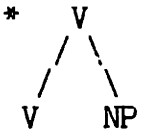
a.



b.



c. *

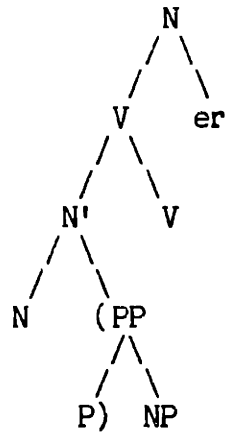


Proponents of a separate autonomous lexical component might think that they have an easy answer to this question: they would simply claim that the relevant compounds are formed in the lexicon, where phrase building operations are non-existent: No phrase building operations, no phrases. I do not believe the answer can be that simple, however, for two reasons. The first reason has to do with examples from English of the following form:

(82) Attila the Hun hater, Jack in the Box admirer, Jack the Ripper chaser

These examples are perhaps all a little unusual, but I think that they are passable. Their structure is presumably all something like:

(83)



All of these, then, contain phrases, which would argue against the position that the reason no maximal phrases occur as sisters to V in synthetic compounds is that synthetic compounds are formed "in the lexicon." Of course, one could fairly object that all of the examples I gave are highly lexicalized phrases and that they therefore must be listed in the lexicon anyway; as phrasal entities, therefore, they would be exceptional. However, notice that the same would seem to be true of names such as The Bronx or The Hague and yet, as we have seen, the definite article in those NPs does not occur in synthetic compounds. What does this mean? What it seems to mean is that the relevant constraint is that the sister to V cannot be an X-max

One may well wonder why one does not find more compounds of the form of those in (82). For instance, why are the following odd?

(84) ??dog in a hat catcher, ??mice with lice chaser, ??cat with a black spot fancier

I suspect that this has to do with a separate "adjacency constraint" to the effect that in the normal case, heads in a construction such as synthetic compounds have to be adjacent. So, dog and catch in ??dog in a hat catcher are not adjacent and therefore the construction is odd. In the case of Jack in the Box admirer, Jack in the Box, even though it is clearly phrasal in construction, is still somehow the relevant semantic head (of itself) and therefore may count as the head for the purposes of this constraint. Needless to say, there are a few ideas which need to be firmed up here, but there is some evidence for such a constraint independent of these particular examples. Fabb (1984), notes that AP complements are impossible in attributive adjective constructions in English: so, *a capable of every sacrifice friend is bad. In German, however, as Fabb notes, the complements in the AP may precede the adjective, which also precedes the modified noun. And in German the equivalent of the bad English example just given is perfectly grammatical:

(85)

ein jedes Opfer faehiger Freund
a of-every sacrifice capable friend
"a friend capable of every sacrifice" (Fabb, p. 131)

This would predict that in languages such as German or Dutch, where complements of nouns and adjectives may precede their heads, Noun constructions containing full NPs should be perfectly grammatical in synthetic compounds since the head of the N' could nevertheless be adjacent to the verb. This prediction is borne out by the following data from Dutch

(thanks to Laszlo Maracz for providing these data):

(86)

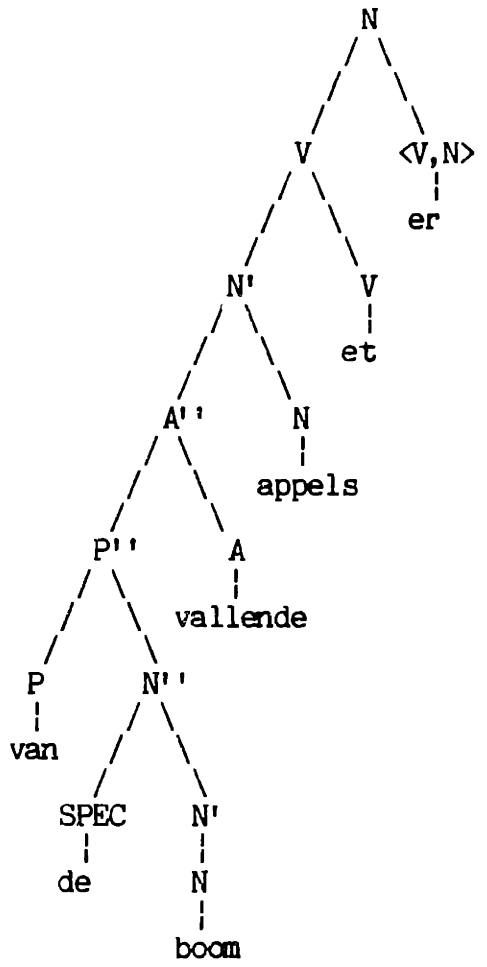
a. de van de boom vallende appels et -er
the from the tree falling apples eat-er
"the eater of apples falling from the tree"

b. de in de oven gebakken cake et -er
the in the oven baked cake eat-er
"the eater of cakes baked in the oven"

c. de in de tuin groeinde tomaten bewonder-aar
the in the garden growing tomatoes admire -er
"the admirer of tomatoes growing in the garden"

The structure of the Agent N in (85'a) is:

(87)



But this clearly contains a maximal projection[8].

Another piece of evidence that the sister of V in synthetic compounds is

---

8. It is interesting, too, that the two examples we have seen, where head adjacency is required, namely modifier-head constructions and synthetic compounds (on the semantics of which see section 2.3.5.2., below), both are argued to involve thematic discharge via theta identification--i.e. modification. Perhaps such head adjacency is required of theta identification constructions.

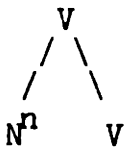
formed in the syntax comes from the fact, already noted, that adjective noun constructions occur within synthetic compounds:

(88) [American History] teacher, [green dog] catcher, [gray elephant] hunter, [blue China] smasher

Unless Stowell (1981) is right in assuming that adjective noun constructions are composed in the lexicon, these constitute further evidence for the syntactic nature of synthetic compounding.

I shall now try to answer the question as to why the nominal element to the left of V in synthetic compounding must be non-maximal. The problem really reduces to the following: Given a structure such as

(89)

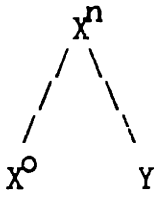


can we deduce the non-maximality of n, that is, that $n<2$? The answer is yes, I believe, as I argue next.

2.3.5.2 Theta Marking within Minimal Projections.

Let us assume that the following principle holds of theta marking (taken in the formal sense of Higginbotham, 1985b, as discussed above):

(90)
In a configuration:

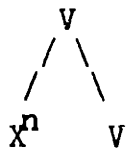


X^o discharges its theta role to Y via theta-marking if and only if n=1.

What this says is that an element discharging its theta role to a complement may only discharge to said complement via theta marking when the element is the head of an increasing sequence of projections, i.e., of a phrase. Furthermore, if a head of a phrase discharges its theta role to a complement, then it must do so via theta marking. So, a verb will theta-mark its object, and a N may theta mark its complement within an NP, but an adjective may not discharge its theta role to the noun it modifies via theta marking.

Given this assumption, it will follow that the sister of V within synthetic compounds may not be a maximal projection. First of all, since V is dominated by another node of the same category (namely V) it will follow that the Projection Principle will apply to a construction of the following form (see section 2.3.3. for discussion):

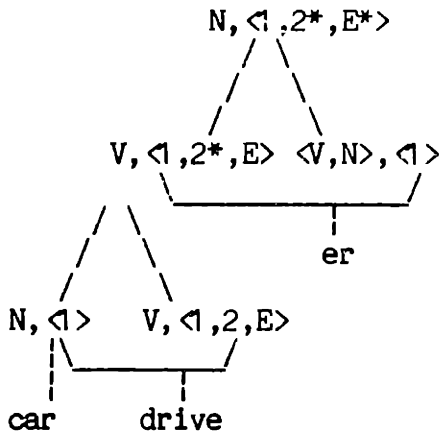
(9')



The Projection Principle will require that all internal arguments of V must be satisfied within this projection. How can the internal theta roles be discharged? By the assumption we have just made, they cannot be discharged by the V theta marking the X^n. Theta binding is presumably out since there is no appropriate binder. That leaves theta identification. We will thus have to identify the internal thematic role of the verb with some other thematic role, and since we want to discharge the verb's role to the X^n it follows, from the fact that identification links two open positions, that there will have to be an open position in the thematic grid of the sister. But it will then follow that the sister could not be a maximal projection since in that case there would be no open position (since all theta positions are discharged within a nominal maximal projection).

The structure that I am proposing for the whole synthetic compound will thus look like the following:

(92)



The * on the 2 at the top V node indicates, as usual, that that theta role has been discharged, this time under identification. The semantics of this form should be:

(93) v(x,N) <<=>> Actor(x) & (Ey)(Ee)[car(y) & drive (x,y, e)]

This seems to be the correct reading[9].

Note however that (92) is not absolutely correct. We have to insist, I

---

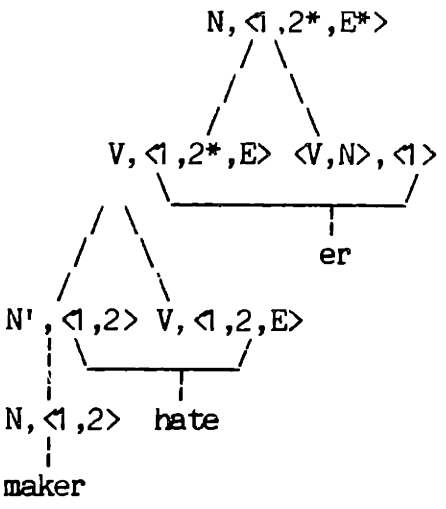
9. Still, for some speakers N-V constructions such as flower arrange may be transitive:

(i) I flower-arranged the tulips.

Assuming that the relation between flower and arrange here is identical with that in the corresponding synthetic compound flower-arranger, the question arises as to where the internal theta-role really is closed off. It's clearly closed off for the synthetic compound as a whole (*he's a cookie-maker of oreos), but it's possible that it is not closed off at V as suggested above. Still, examples like (i) are not generally agreed to be good, so the issue is not clear.

believe, that N project to N' in any case in these constructions. In that case, since N now itself projects to a dominating node of the same category, its own theta grid must be satisfied. We could then say that the internal theta role of the verb and the theta role of the noun discharge to each other via identification. This will mean that the noun in effect modifies the position in the verb's grid, and that simultaneously the verb discharges its theta-role under identification to the noun. This point about modification is crucial since we would expect similar behavior between synthetic compounding and adjective noun modification; one such similiarity has already been found, namely the head adjacency requirment discussed in section 2.3.5.1. The important consequence of this added assumption that N projects to N' is the following: relational nouns (such as other -er nominals) will have to, by the Projection Principle, somehow have their thematic requirements satisfied within the N'; given that this is impossible in many cases for other reasons (c.f., the possible adjacency requirement discussed above), forms such as *maker hater should be bad, as indeed they are. Maker has an internal theta role (derived from the verb make), which must be satisfied. Since the N' does not allow a complement, by the adjacency requirement, it cannot be satisfied in this way, yet it must be satisfied, so the construction is out:

(94)



Contrast this with cookie maker hater, which is fine.

2.3.5.3 On the Application of the Projection Principle within Synthetic Compounds.

The application of the Projection Principle to the V in synthetic compounds has the effect of forcing all of the internal theta roles of the verb to be satisfied within the verbal projection. This has the consequence that constructions of the following form are bad:

(95) *shelf-putter of books, *book-putter on shelves.

Examples of this form have been noted and discussed in Lieber (1983) and Selkirk (1982). Under the current approach these forms will be out for the following reason: put has two obligatory internal theta roles. Since the Projection Principle applies to the $[_V$ N V] constituent of the compounds it

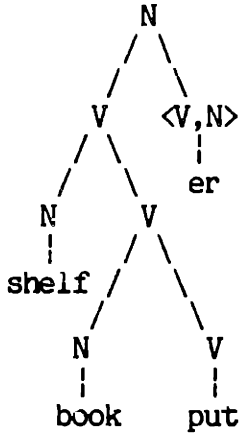
will be necessary for the verb to discharge the two theta roles within the compound. It cannot discharge two theta roles to one argument, because of the Theta Criterion, yet in both of the examples in (95) there is but one nominal preceding the verb. These forms should thus be out since one or other of the two theta roles will be left undischarged.

(This system of principles and assumptions has much the effect of Selkirk's First Order Projection Condition (Selkirk, 1982). She writes:

> In fact it seems appropriate to see these restrictions on compounds as instances of a more general condition on the satisfaction of argument structure in syntactic representation. I will state the condition in (2.41) and suggest that <u>first order projection</u> be defined as in (2 42):
>
> (2.41) <u>The First Order Projection Condition (FOPC)</u>
>
> All non-SUBJ arguments of a lexical category X_i must be satisfied within the first order projection of X_i.
>
> (2.42) The <u>first order projection</u> (FOP) of a category X^n_i is the category X^m_j that immediately dominates X^n in syntactic representation (i.e., in either S[entence]-syntactic or W[ord]-syntactic structure).)

One may well ask why the following construction is not possible:

(96)



That is, why is only one nominal allowed in such a construction? I shall suggest that the answer is Case, and I discuss this next.

2.3.5.4 Case Assignment in Synthetic Compounds

I would like to suggest that, in addition to theta role discharge which is obviously relevant for analyzing synthetic compounds, (structural?) Case assignment is also relevant; that is, the verbal head will assign Case to its nominal complement. So in car driver, drive will assign Case to its nominal complement car. Assuming that Case assignment is only possible under adjacency (see Stowell, 1981), it will follow that a verb could only Case mark one complement. Therefore a form such as *shelf book putter would be out simply because shelf could not get Case[10].

---

10. One might wonder why overt morphological case is not found ir compounds (not in English, of course, but in other languages). The answer to this

One thing which is striking about this assumption is that, if it correct, then Case marking will be to the left in these constructions (at least when the ordering is imposed in the PF component), unlike the phrasal assignment of Case in English, which is to the right. The same point can be made for the verb's thematic discharge which is also to the right in phrases in English, but to the left in synthetic compounds. Travis (1984) notes this fact and states (p. 106):

> In English we see an example where only head-initial/final need be specified and the direction of θ-role assignment will follow. The VP is head-initial so complements follow the verb (150). X^0s, however, are head-final, so complements must precede the verb.
>
> (150) keep books
> (151) bookkeeping, bookkeeper

Travis only mentions theta-role assignment, but we can as easily extend the analysis to Case. This asymmetry between phrases and X^0s in English is undoubtedly a historical relic of a time when English, like its West Germanic relatives Dutch and German, was head-final. In Chichewa, as we shall see below, synthetic compounds follow the ordering of the phrases and are thus head-initial (Chichewa is SVO)[11].

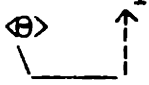
---

might be that such overt Case morphology is limited to marking Case assigned to NPs.

11. Still, there is another way to look at this which has nothing to do with Case directionality. As we have noted, we are analyzing synthetic compounds as cases of modification; the N' modifies the V by identifying

One interesting consequence of the assumption that Case is involved in synthetic compounding is that there are no synthetic compounds formed with unaccusative verbs. Unaccusatives (c.f. Perlmutter, 1978) such as arrive have standardly been assumed within the GB literature, following Burzio (1981) (who calls them ergatives) to assign internal theta roles but no external theta roles on the one hand, and no Case on the other. This lack of Case assignment forces the object NP to move to find Case, and it does so to the subject position:

(97) $John_i$ [arrived t_i]



So the surface structure of a sentence with an unaccusative verb ends up looking like the surface structure of a sentence with a true intransitive--i.e., a verb like walk which actually assigns an external thematic role.

We cannot test the lack of synthetic compounds with unaccusatives using -er forms since, because of the lack of external argument, there would be

---

its theta-role with a position in the V's grid. Now, in English, modifiers canonically precede heads, which would derive the order N-V. In Chichewa, modifiers come after heads, which would derive the order V-N. The latter is also true of Romance, where we find constructions such as French caisse-tout ('break-all', i.e. klutz) or Spanish tocadiscos ('play-disks,' i.e. disk-jockey.

nothing for the -er's own thematic role to link to. However we can show it with -ing nominals. These forms have many of the same properties as -er compounds as the following examples illustrate:

(98)
a. dog catching (c.f. dog catcher)
b. green dog catching (c.f. green dog catcher)
c. *shelf putting of books (c.f. *shelf putter of books)

It seems as though, therefore, the -ing compounds have the verb's internal thematic role assigned to the sister of the verb just as in the -er compounds. However, examples such as the following are ill-formed:

(99) *man-arriving, *ship-sinking (out as 'The ship sinks ' not as 'Somebody sinks the ship'), *star rising

This makes sense if we assume that Case assignment is relevant in synthetic compounds since a verb like arrive will not have a Case to assign to its sister.

One class of cases which is interesting are compounds where the sister to V expresses a location or direction:

(100) church-goer, theater-goer, wall-sitter...

These are not derived from verbs such as *church-go, *theater-go, *wall-sit and are thus like regular -er compounds in this respect. It seems in fact that they involve the regular assignment of a thematic role within the V-projection. Yet the internal theta role of verbs like go and sit is assigned in sentences to NPs governed by directional or locational

prepositions:

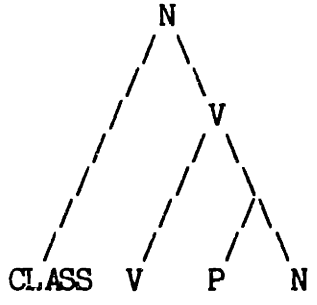
(101) go to church, go to the theater sit on the wall...

I do not have anything particularly interesting to say about these forms. It may well be the case that verbs such as sit and go exceptionally allow Case assignment within synthetic compounds. This would be consistent with the fact that not all such verbs allow compounds of these forms:

(102) *church-comer, *bed-lier, ??floor-stander....

So sit and go may well be marked. Also, it is interesting that in Chichewa, as I shall describe in more detail below, similar verbs (such as lie and sit) allow preposition-like elements within the synthetic compound, apparently to mark case yielding forms with the following structure:

(103)



We have presented an account of synthetic -er compounding. This is far from a discussion of everything interesting which there might be to say about such forms. Nevertheless, I think that I have succeeded in showing that the core properties of these constructions follow fairly naturally

from reasonable principles of grammar, in particular, without recourse to the assumption that synthetic compounding is done in the lexicon. I now turn to a comparison of this approach with that of Lieber (1983), the most recent and complete treatment of synthetic compounds in the GB literature.

2.3.6 A short history of the study of Synthetic Compounding and a comparison with Lieber (1983).

One thing which has been observed about the element which compounds with the V in synthetic compounds is that it is the element which occurs immediately to the right of the verb in the corresponding sentence. In terms of our story, of course, this makes sense insofar as the first sister is generally the subcategorized and Case-marked complement of the verb in the sentence, and given the analysis presented in the last section, this, and only this, will be licit as a sister of V in the synthetic compound.

Roeper and Siegel (1978) were the first to formulate a principle describing this state of affairs. They called it the First Sister Principle:

(104)

First Sister (FS) Principle

All verbal compounds are formed by incorporation of a word in the first sister position of the verb.

The mechanism that they propose for doing this is the Compound Rule,

formulated as follows:

$$[[\text{empty}] + \text{verb} + \text{affix}]\ [_{X[+N]} + \text{word}\ ]\ W \Rightarrow$$
$$\quad 1 \qquad\quad 2 \qquad 3 \qquad\qquad\quad 4 \qquad 5$$

$$[[+\ \text{word}] + \text{verb} + \text{affix}]\ \ W$$
$$\quad 4 \qquad\quad 2 \qquad\quad 3 \quad 0\ 5$$

What this is is a transformation (a "lexical transformation", to use Roeper and Siegel's terminology, originally due, according to them, to Vergnaud) which takes a [+N] element generated in the first-sister position after the verb, and moves it into a slot provided in the deverbal compound to its left. For example (a) will be transformed into (b):

(105)

(a) $[[\text{empty}] + \text{drive} + \text{er}\ ]\ [_{N}\ \text{truck}]\ W \Rightarrow$

(b) $[[\text{truck}] + \text{drive} + \text{er}]\ W$

Lieber (1983) presents arguments against Roeper and Siegel's story. In particular, as she points out, their analysis makes use of some very powerful mechanisms, including the lexical transformation stated above. Such mechanisms have come to be eschewed in Transformational Grammar, and with good reason; their effects are often reducible to principle which themselves can be argued to have fairly general application, thus yielding a more plausible and explanatory system.

Lieber provides an analysis of English synthetic compound which makes use of a set of principles capable of making the right predictions. One of

these principles is her Argument-Linking Principle:

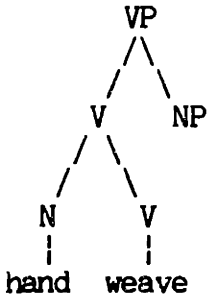
(106) Argument-Linking Principle

a. In the configuration $[\ ]_{\{V\ P\}}\ [\ ]_{\alpha}$ or $[\ ]_{\alpha}\ [\ \ ]_{\{V\ P\}}$, where α ranges over all categories, {V P} must be able to link all internal arguments.

b. If a stem $[\ ]_{\alpha}$ is free in a compound which also contains an argument-taking stem, α must be interpretable as a semantic argument of the argument-taking stem, i.e. as a Locative, Manner, Agentive, Instrumental, or Benefactive Argument.

In a case like hand weave, the verb weave passes its argument structure onto the entire compound, and since this compound is itself the head of a phrase, it satisfies its argument structure outside the compound; hand is therefore free in the compound and must be interpretable as a semantic argument, in this case an instrumental

(107)



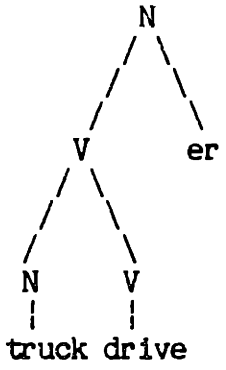
Turning now to synthetic compounds, we find that the left-hand member can be assigned a thematic role by the verb. So, in truck-driver, truck satisfies the internal O-role of drive. Why doesn't driver necessarily satisfy its argument structure outside? Lieber suggests the following (p.

268):

> ...since it [the verb, RS] is contained within a noun, its argument structure cannot percolate to any node higher in the compound tree. (It is assumed that features of one category cannot percolate to a node dominated by another category.)

So in truck-driver, argument structure is blocked from percolating across the category change from V to N. The argument structure of drive may thus be satisfied inside the compound:

(108)



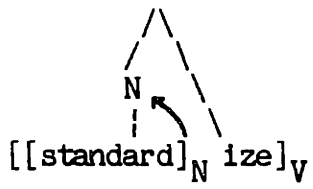
Although Lieber only mentions the fact that category change blocks argument structure percolation as a parenthetical remark, it is crucial for her analysis, since this is the only reason why drive in X-driver would not, like its main-verbal counterpart, obligatorily satisfy its argument structure outside, if it satisfied it at all.

On closer reflection, though, it is not at all obvious that the restriction is motivated. It is certainly not motivated by Lieber s own Feature Percolation Conventions:

(109) Feature Percolation Conventions (Lieber, 1980)

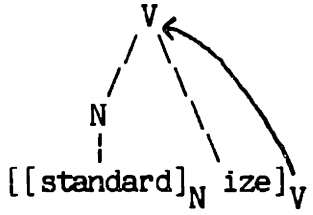
a. Convention I

All features of a stem morpheme, including category features, percolate to the first non-branching node dominating that morpheme, for example:



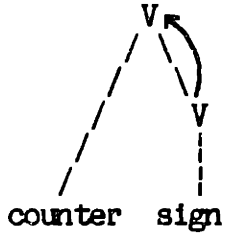
b. Convention II

All features of an affix morpheme, including category features, percolate to the first branching node dominating that morpheme. For example:



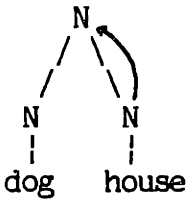
c. Convention III

If a branching node fails to obtain features by Convention II, features from the next lowest labeled node automatically percolate up to the unlabeled branching node. For example:



d. Convention IV.

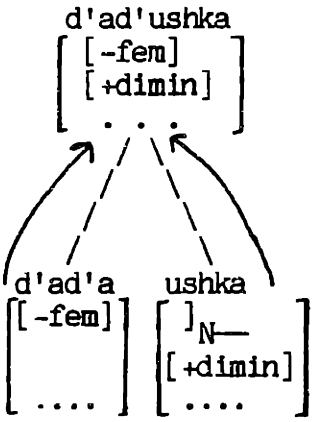
If two stems are sisters (i.e., they form a compound), features from the right-hand stem percolate up to the branching node dominating the stems. For example:



It is convention III which is of interest to us. This could be interpreted to mean that just in case a branching node fails to get any features at all via Convention II, then and only then will features from the next lowest node percolate to the mother node. While this is a possible interpretation, apparently this isn't the intended one as evidenced by many examples from Lieber (1980). The following example from Marantz (1984a) serves to illustrate the point nicely. In Russian there is a diminutive suffix -ushka which attaches to nouns ending in -a of either masculine or feminine gender. The interesting characteristic of this suffix is that it does not change the gender of the noun to which it attaches. So, where d'ad'a 'uncle' is masculine, so is d'ad'ushka 'uncle dimin.', whereas with baba 'grandmother', which is feminine, we have babushka which is also feminine. Now, -ushka presumably does have information of its own, such as the fact that it forms diminutives, and this information would presumable percolate via Convention II. It does not, however, carry features for gender, and these must percolate, via Convention III. We therefore get a

representation such as:

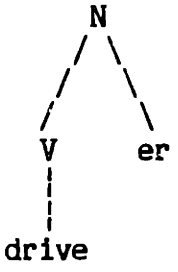
(110)



The intended interpretation of Convention III, then, is that a branching node gets features from the next lowest node whenever those features are not filled in via percolation from the affix.

Turning now to the synthetic compounds, it is not easy to see, given Convention III, why the argument structure should not percolate up from the V to the N in a structure such as:

(111)



Under Lieber's analysis, at least, -er does not carry an argument structure of its own. Therefore, the argument structure of drive ought to be free to percolate. It is not obvious what principle (other than mere stipulation) would rule it out.

In fact, even if the percolation conventions were to rule out percolation from a V node to a dominating N node, this would not be a desirable result since there are cases, as we have seen, where it is necessary to argue that the verbal theta grid is inherited by the noun: the most telling examples are the ones like *maker vs. cookie maker or maker of cookies as discussed in section 2.3.2. Lieber's analysis predicts, apparently, that maker of cookies should be bad simply because, given the fact that theta roles for her do not percolate across categories, there should be no way in which the internal argument of make could be satisfied outside and further how cookies could therefore get a theta role in the first place. It should be noted, too that Lieber's analysis wrongly predicts that verbal theta roles cannot be assigned by derived nominals--on which see section 2.4.

In the analysis presented in the preceding sections of synthetic compounding, we showed how an interpretation of the Projection Principle could derive the fact that if the V is present in the relevant configuration, then all of its internal thematic structure must be discharged within the compound. This is part (a) of Lieber's Argument

Linking principle, and this can now be seen to reduce to the Projection Principle and the Theta Criterion. (Part (b) I essentially agree with; that is the part which, among other things, governs interpretation in Root compounds such as dog house.) Also, since we do not stipulate that thematic structure may not percolate across categories, we are able to account for the large class of cases where Agent nominals discharge their theta roles outside the compound.

However, if deverbal nouns may in general inherit the argument structure from the verb, and if this argument structure may be satisfied outside--and both of these seem to be correct generalizations, then we have lost the explanation that Lieber has for the existence of synthetic compounds versus the non-existence of N-V compounds where the verb's argument structure is satisfied within the compound; note the non-existence of such (verbal) compounds as *dog-kill, *deer-shoot, baby-see:

(112)

a. *John dog-kills. (i.e. kills dogs)
b. *Mary deer-shot. (i.e., shot a deer)
c. *Rachel baby-liked. (i.e., liked a baby)[12]

The ungrammaticality of these will follow from Lieber's formulation since the argument structure of the verb must be satisfied outside: again, this

---

12. Some speakers accept compounds such as flower arrange, as Lieber points out.

is because the V heads a VP and the theta-grid of the V thus must percolate to the VP and be satisfied within that domain, rather than merely within the domain of the compound. This will contrast with the synthetic compounds where the argument structure may be satsified inside:

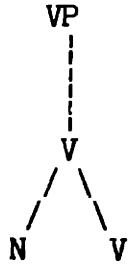
(113)

a. John is a dog-killer.
b. Mary is a deer-shooter.
c. Rachel is a baby-liker.

Yet if argument structure can be satisfied outside derived nouns, as we have shown, there can be no blocking of argument structure percolation by category changes, and the difference between verbs and deverbal nouns disappears in this regard. Therefore, it seems that there must be some reason why, for all speakers, sentences such as those in (112) above are bad.

So why are constructions of the following form bad?

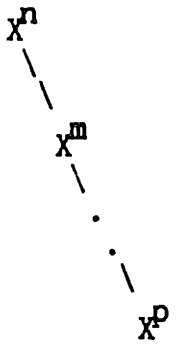
(114)



where N picks up (one of) the internal theta role(s) of V? It could in fact be that a principle very similar to Lieber's conception of the

operation of the percolation conventions in such instances is the appropriate idea. We might assume for instance, that given a path of nodes in a tree of the form:

(115)



where n>=m>=...>=p, then the thematic role must be discharged within the maximal allowable i, where i is an element of {n, m,..., p}, this of course being subject to other constraints such as Case assignment and so forth. We might then insist, following this idea, that a V which is in a chain of V projections up to VP, actually discharge its theta role at the level of the VP, that is via theta marking to a phrasal complement. The ideas here obviously need to be worked out more formally, but I believe them to be essentially on the right track. Note also that synthetic compounds are consistent with this idea in that the verb's theta-role is discharged within the maximal projection of V, which is in this case sister to _-er_.

What I have argued, then, in this brief discussion of Lieber's analysis of synthetic compounding in English is that my own analysis is able to

handle her data and does not suffer from the drawbacks of hers, in particular the incorrect stipulation that theta grids may not percolate across categories. I would now like to turn to a quick discussion of Agent nominal compounding in Chichewa, which differs in some interesting ways from the corresponding English constructions.

2.3.7 A short note on Chichewa synthetic compounds.[13]

Chichewa, a Bantu language spoken in Malawi, has synthetic compounds which resemble in many ways those of English, but with some interesting differences as we shall note. In particular, prepositions are allowed as Case-markers for the complement of V in synthetic compounds in Chichewa; furthermore, as we shall see, assignment of the theta role outside the derived noun is not possible in Chichewa, and there is thus no Chichewa equivalent of <u>maker of cookies</u>. Below, I give some examples of such compounds:

(116)

a. m- pala- matabwa 'carpenter'
   Class scrape wood

b. chi- pha- dzuwa 'beautiful woman'
   Class kill sun

c. m- khala pa mwala 'rock sitter'
   Class sit on rock

---

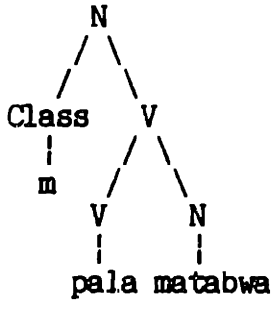
13. I wish to thank Sam Mchombo, who is a native speaker of the Chinyanja dialect, for providing most of the data in this section.

d. chi- lima- mpunga 'rice cultivator'
Class dig rice

The prefixes m- and chi- are regular noun class markers in Chichewa. M- is the (singular) prefix for Class I, the so called 'personal' class (Watkins, 1937) into which fall a large number of nouns referring to people. Chi- is the (singular) prefix for Class VI, which contains nouns often referring to non-animate things. Some of these forms are highly metaphoric and hence lexicalized, such as the (b) example. On the other hand, forms like (c) show this to be a very productive process since Mchombo informed me that this particular example was coined by him and a friend of his to refer to an individual whom they regularly saw sitting on a certain rock on the side of a path.

I shall assume that the structure of these is as in (117) below We will not concern ourselves at this point with the role of the class marker; I return to a discussion of this below and concentrate for the time being on the properties of the constituent headed by the top V:

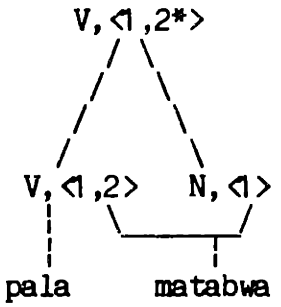
(117)



'wood scraper' (= carpenter)

This is more less the structure given for these constructions by Mchombo (1978). The Projection Principle, given that this is the right structure, will therefore require that the verb's internal arguments be satisfied within the compound. By the suggestion given in section 2.3.4.2. above, it will follow that the verb will have to discharge its theta role via identification. Therefore the entire structure for the V constituent will be:

(118)



Therefore, as in English, as also argued in 2.3.4.2. the sister of the V cannot be a maximal projection. This is correct since in fact, Chichewa seems to disallow anything but a plain noun here, as the following example indicates:

(119)

??m- kusa agalu akuda
Class catch dog black
'black dog catcher'

Furthermore, it should not be possible to have forms analogous to *child-giver of candy or *candy-giver to children and this is also correct:

(120)

a. *m- patsa-ana masiwiti
Class hand child sweets
'a child-hander of sweets'

b. *m- pereka-siwiti kwa ana
Class hand sweets to child
'a sweet-hander to children'

That Case-assignment is relevant in these constructions is indicated by two facts. First of all, there can, as in English, be at most one N within the V constituent even when thematic requirements would dictate that two nominals must be present:

(121)
*m- patsa-ana- siwiti
Class hand child sweets

We argued that the absence of such constructions in English should be reduced to Case assignment.

The other indication that Case is involved has to do with intransitive verbs which allow their nominal complements to be Case-marked with a preposition such as pa or m (following the analysis of Mchombo (1978).)

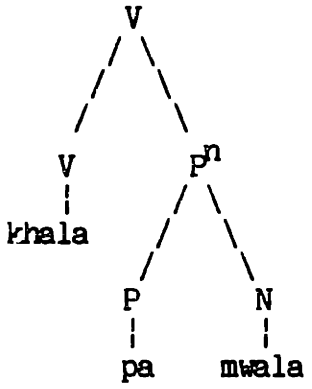
(122)
a. m- khala pa mwala 'rock sitter'
Class sit on rock

b. chi- gona m- bawa 'drunkard'
Class sleep in bar

c. chi- lima pa dzala 'dump digger'
Class dig on dump

Following Mchombo's own analysis, we will say that these have the following structure:

(123)



'rock sitter'

Here, P^n is some projection of P, though not necessarily a P^{max}; in any event, we may assume that the status of P^n is not particularly important insofar as the apparent function of the prepositional element in these constructions is to assign Case where the verb fails to do so on its own. Apparently, no such mechanism is available in English.

Unaccusatives cannot be tested since there is no synthetic compounding in Chichewa equivalent to English <u>-ing</u> compounds.

Finally, in Chichewa, unlike English, Case and theta-role assignment is rightwards within synthetic compounds just as it is in phrases; for example:

(124)
mbuzi zi- na- mwany-a chikho
goats subj-past-smash-indic gourd
'The goats smashed the gourd.'

We have examined a few of the properties of the V-constituent in Chichewa synthetic compounds, and pointed out the similarities and the differences between those and the English equivalents. On the whole, it seems that the constructions are quite similar in both languages, and since these languages are obviously completely unrelated and furthermore completely at variance with one another when it comes to the syntax at large[14], it seems that whatever principles one brings to bear upon these sorts of constructions must be very basic and widespread in application.

We now turn to the Class prefixes and their role in the constructions. The prefixes chi- and m- really are general noun forming prefixes (the examples are transliterated from the lists in Watkins (1937)):

(125)
Class I: muntu 'person'; mchewa 'Chewa person'; mzamba 'midwife'; mnasi 'friend, neighbor'.

Class VI: chichewa 'Chichewa language'; chinangwa 'cassava'; chiwale 'coconut tree'; chitanda 'corpse'

Now, it seems reasonable to suggest that these prefixes have argument positions of their own which are then percolated to the noun dominating

---

14. In particular, Chichewa, unlike English, has a complex morphology for marking grammatical function changing; see Marantz (1984a) and also Baker (1985c).

them; it also seems reasonable that there might be some restrictions which generally apply to the referents of nouns created by these affixes. The entries thus might look like:

(126)

a. $M' = \langle M_{\langle O,N \rangle} \langle 1 \rangle, mu \rangle$[15]
   pre(mu)
   Human(1)

b. $CHI' = \langle CHI_{\langle O,N \rangle} \langle 1 \rangle, chi \rangle$
   pre(chi)
   Non-human-thing(1)

So m- and chi- are affixes which attach to other elements and make nouns. They have a single argument position which is specified as being human or non-human-thing respectively. They are therefore simply nominalizers and carry no special semantics with them other than the reference to human versus nonhuman.

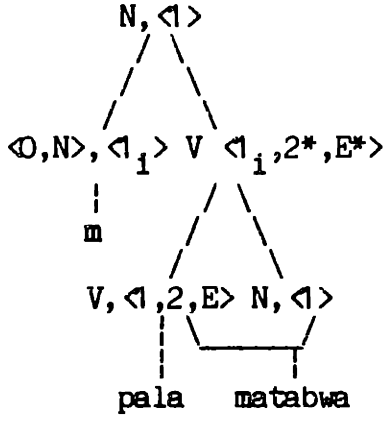
What does not seem appropriate is to suggest that, like English -er, m- and chi- have the alternate specification that the 1 may be an Actor. In English -er is used productively only in producing Agent nominals. The Chichewa prefixes, on the other hand, are used for a whole variety of nouns, the only restrictions which generally apply being that nouns in m- are human and nouns in chi- are not and generally refer to things.

---

15. Mu is the actual phonological form; see Watkins, p. 27.

Say, then, that the prefixes are not specified as Actors and are therefore not marked, as is English _-er_, to link up to the external theta role of the verb. Again, all these prefixes are are nominalizers; they do not occupy a special place in Chichewa grammar like English _-er_. In the examples we have seen, all of the internal theta roles have been discharged, and we may assume that the Event is discharged, as in English, via binding, although there is a slight trick here since we had assumed for English that the Event position was discharged in the transition from V to N. The external theta role need not be discharged at all, in fact, and it may merely be indexed with the prefix:

(127)



'wood scraper' (= carpenter)

We have not developed a semantics for theta indexing, in part because of the fact that it was introduced for capturing theta relations in more or less untransparent morphology. Nevertheless, I suggest that the following be taken to be the semantic representation of (127):

(128) v(x,N) <<=>> Human(x) & [E(y) E(e)[wood(z) & scrap(y, z, e)]] & x=y

So indexing equates some object with a particular position in a theta grid, although it does not discharge that position. Remember that we have assumed that the External thematic role need not be discharged.

This set of assumptions for Chichewa synthetic compounds, if it is right has an interesting prediction. Given that the noun prefixes m- and chi- are not specified to identify with (that is discharge their thematic role under identification to) the external (Actor) theta role of the verb, it should follow that a verb's thematic roles will not need to percolate across the categorial break from V to N (see the Cross-categorial Theta-grid Percolation Convention, section 2.3.3.) In fact, since the affix, the syntactic head of the word has its own theta grid to percolate, it will follow that the verb's cannot percolate. This predicts that NPs of the form driver of the car will be ungrammatical in Chichewa with these synthetic compounds since no thematic grid is inherited on the noun from the verb. This seems to be correct since forms equivalent to the English examples are non-existent in Chichewa, according to Mchombo (p c.). There is another form which looks similar to these, but which is formed with the relativizing prefixes u+a (==>wo-). These clearly contrast with the synthetic compounds, in that the recipient of the internal theta role is phrasal; one way to test this is with referentiality:

(129)

a. *m- dya makoko$_i$ si- a- wa$_i$- kanda ndi jamu
Class eat husks not subj obj(them) like with jam
*'A husk$_i$-eater doesn't like them$_i$ with jam.'

a. wo- dya makoko$_i$ si- a- wa$_i$- kanda ndi jamu
rel eat husks not subj obj(them) like with jam
'An eater of husks$_i$ doesn't like them$_i$ with jam.'

As has been observed for a long time now (c.f. Postal 1969), words are anaphoric islands (see also, Chapter 3 of this thesis.) The fact that (129b) is grammatical is evidence that in this example makoko outside the word wodya. On the other hand (129a) clearly shows that mdya-makoko is one word, and in fact must be one word, confirming the prediction made above that synthetic compounds in m- and chi- cannot assign the verb's thematic roles outside.

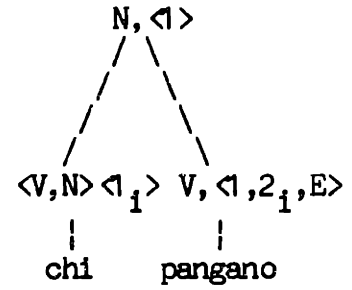
Of course, given that m- and chi- can index a thematic role without discharging it, we would expect that they could form deverbal nouns which "pick up" one of the internal theta roles of the verb. This appears to be correct:

(130)

tseka 'shut'; chi-tseko 'door'
pangana 'promise'; chi-pangano 'promise'
sonkhana 'get together'; msonkhana 'meeting'

The structure of one of these is given as follows:

(131)



The verb's other theta roles can of course, just sit there, locked up in the morphology, as we saw in the English examples discussed in section 2.3.3.

2.4 The Syntax and Semantics of Derived Nominals.

We now turn to a discussion of one of the more popular topics in the literature, namely the behavior of derived nominals. This is another example of a highly productive process which will, as I shall show, demonstrate the viability of an analysis of word-formation whereby syntactic and semantic well-formedness is derivable from general principles of syntax and semantics rather than specifically lexical principles. A typical example is the following:

(132)

a. Eric destroyed the jungle fort.

b. Eric's destruction of the jungle fort was a surprise[16].

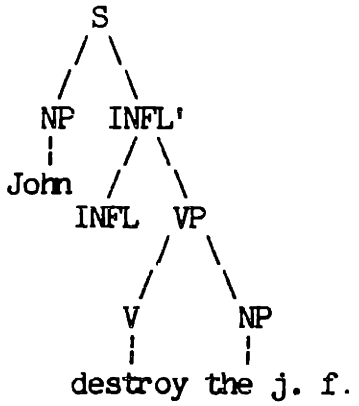
In (132a) destroy assigns the internal theta role of Patient (= destroyee) to the NP the jungle fort; the entire VP assigns, via predication, the external theta role of Agent (= destroyer) to Eric. Similarly, in (132b), the NP the jungle fort is assigned the theta role of Patient, and somehow or other, though presumably not via predication, Eric picks up the external theta role of Agent. Why not via predication? Because predication is canonically assumed to be assignable from maximal projections, and there is no maximal projection available here. Contrast (133a) and (133b):

---

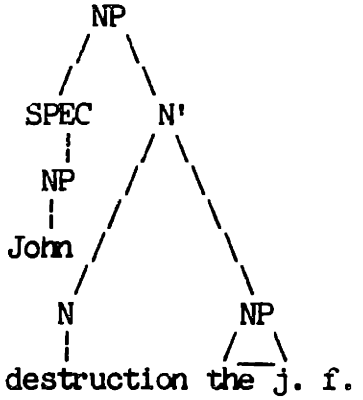
16. This particular example is interesting in that it brings up the point of intensionality, familiar from sentential complements in belief contexts, for example. Derived nominals can be opaque contexts: If Clark Kent's destruction of the Empire State Building was a surprise to me, it does not follow that Superman's destruction of the Empire State Building was a surprise to me since I might not know that Clark Kent is Superman. Derived nominals can also be intensionally interpreted in odd ways: in particular, The bomb's explosion was a surprise can mean that the fact that the bomb exploded was surprising or alternately the way in which it exploded was surprising. The latter interpretation is not available in It was surprising that the bomb exploded. I will not deal further with this issue here.

(133)

a.



b.



Since the N' is not a maximal projection it cannot, under standard assumptions, be a predicate of the subject of the NP (see Rothstein, 1984; Williams, 1980; Higginbotham, 1985b.)[17] There are semantic reasons too for believing that predication does not take place in NPs; as Higginbotham (1984) points out, there is no reason, given that N' could be predicated of the subject of the NP, why John's dog could not mean that John is a dog.

---

17. For convincing arguments on the internal structure of NP, in particular on the existence of N', see Giorgi (1984).

It clearly does not mean this. Similarly Eric's destruction of the jungle fort does not mean that Eric destroyed the jungle fort. Rather it refers to the action, or, as I shall suggest, the Event, of Eric's destroying the jungle fort.

In this work I shall be concerned primarily with a somewhat different question than what has interested researchers on this topic in the past. I shall be investigating the morphological relationship between derived nominals and the verbs from which they are formed. However, I shall also be considering the syntactic behavior of these forms, but mainly with an eye to uncovering their morphological structure.

2.4.1 The Secret of NOM: the Morphological relationship between Verbs and Derived Nominals.

Let us assume, to begin with, that there is an abstract morpheme--call it NOM--which is responsible for nominalizing verbs. It will have various phonological spell outs depending upon the particular verb. Here are some examples:

(134)

a. Eric's DESTRUCT-NOM of the the jungle fort
   destruction

b. John's EAT-NOM of the rutabagas
   eating

c. Monty's SOLVE-NOM of the problem in Lie Algebras
   solving

d. The hood's EVADE-NOM of the CIA
evasion

e. The engineers DRAIN-NOM of the water from the pond
drainage

In many cases NOM will be spelled out as -tion in many others as -ing and in still others as some other affix such as -age. Note that the -ing nominals have often been called "action nominals" as distinct from derived nominals; there seems to be little reason to separate them since both seem to involve a nominalization which would, in current jargon, be called "lexical" (see Baker, 1985b, for instance for a discussion of the difference between the lexical action nominals and the syntactic gerunds.) I shall henceforth assume that they are part of the same process.

What does the lexical entry of NOM look like? Obviously its phonological entry, given what we have said, cannot be a constant, but will depend upon the particular verb to which the affix is attached; there is nothing unusual in this, of course, since in many languages there are sets of verbal or nominal paradigms where a particular Case or tense affix is spelled out differently according to the particular paradigm chosen. We will therefore concentrate on the syntactic and semantic parts of the entry. I suggest that these look as follows:

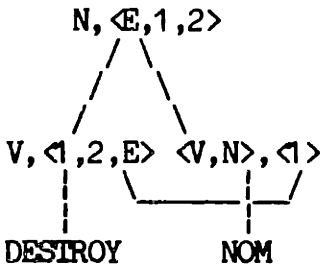
(135)

$$\text{NOM}' = \langle \text{NOM}_{\langle V,N \rangle} \langle 1 \rangle, \ldots . \rangle$$
$$\text{Event}(1)$$

NOM is therefore an affix which attaches to verbs and makes nouns.

Furthermore, like _-er_, it has a role of its own, which in this case is marked as referring to an event. Following the discussion in Section 2.3.2., we will assume that the affix's theta role is discharged under identification to the appropriate position in the thematic grid of the verb, in this case the Event position. We suggest again that since NOM is the syntactic head of the construction, it determines the particular verbal theta role which the nominal picks up as being the one to which it refers. However, since the verbal theta roles remain undischarged, and since there is a productive process of thematic discharge relating one of the positions in the grid with the syntactic head of the construction, then, following the Cross-Categorial Theta-Grid Percolation Convention, the verb's theta roles must be percolated to the dominating N node. In the diagram below, I notate the fact that the Event position is the one to which the N refers by ordering that position first in the grid of the derived nominal:

(136)



The semantics of this construction will be as follows:

(137) v(_<e,x,y>_,_N_) <<=>> Event(_e_) & [destroy(_x_,_y_,_e_)]

That is <e,x,y> is the referent of N just in case e is an event and e is in particular a event of x's destroying y. This seems to be the correct meaning for this noun[18].

But why is the percolation of the Event position licit here? Remember that we said that nouns could not have event positions, but not that they could not refer to events. The latter claim would obviously be false anyway since nouns like event clearly refer to events. What nouns cannot do, as we argued in the section on -er nominals, is have an Event position in addition to referring to something or other. But this is not what is going on here; destruction actually refers to an event of destroying.

The above formulation of the process of nominalizations has a number of co sequences and makes a number of predictions, when coupled with fairly standard assumptions about the structure of NPs, and with the definition of the External Theta Role as given above in section 2.3.2. I discuss these consequences for the remainder of this subsection.

---

18. Interestingly, Johnson (1985) suggests that the following constraint holds of arguments: An argument may bear an AGENT theta-role if and only if it is in subject position of an eventive predicate. Thus, a verb like surround in The trees surrounded the houses, insofar as this does not assign an Agent theta role, also does not have an Event position. This correlates with the fact that an Event nominal is impossible here: *the trees' surrounding of the house.

2.4.1.1 The Derived Nominal's Reference to an Event: Some Consequences.

The first fact that we obviously derive from the representation of NOM as formulated above is that the entire derived nominal should, by definition, be able to refer to an Event (on this, see also Thomason (1985); Barwise and Perry (1983)). This seems to be true as the following examples show:

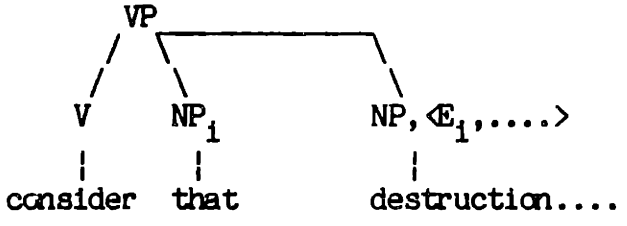
(138)

a. I would consider <u>that</u> (to be) destruction of everything that you have ever tried to achieve.

b. John's destruction of the jungle fort lasted three hours.

In (138a) the NP <u>destruction of everything that you have ever tried to achieve</u> is predicated of the NP <u>that</u>; the sentence is entirely analogous to 'I consider that fat' where <u>fat</u> is a predicate of <u>that</u>. Clearly in (138a) what <u>that</u> refers to itself is an action or Event of some kind. The structure of the relevant part of (138a) would thus be:

(139)



The indexing here marks predication between the Event theta role of the NP and the NP that. Similarly, in (138b), it is clear that the thing which lasted three hours is the event of John's destroying the jungle fort.

A clear prediction that is made by having the derived nominal refer to the Event is that PPs which in sentences modify the Event, such as Instrumentals or Locationals, should be available in such derived nominals. This is correct. Thus, while one cannot say sentences such as (140-142a), as noted in the previous section of this chapter, one can say any of the examples in (140-142b):

(140)
a. *John is a destroyer of jungle forts with a machete[19].
b. John's destruction of the jungle fort with a machete lasted three hours.

(141)
a. *Jones is a butterer of toast with a knives in bathrooms at midnight.
b. Jones' buttering of the toast with a knife in the bathroom at midnight was an unusual event to be sure.

(142)
a. *Those engineers are drainers of ponds with Sweetheart® straws.
b. The engineers' drainage of the pond with Sweetheart® straws was not approved by the Municipal Water Board.

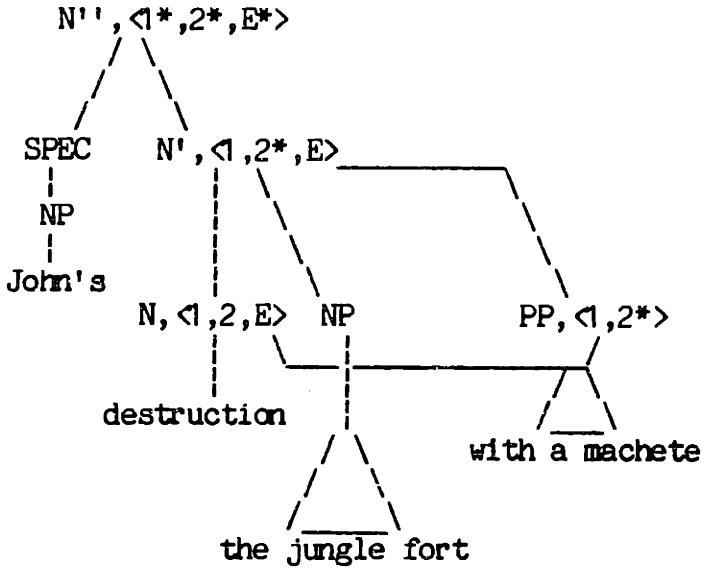
This fact makes perfect sense given that the PP has a thematic role to link

---

19. I.e., the bad reading is the one where John uses a machete to destroy the jungle forts.

to something, and that in these constructions it will link to the theta role associated with the noun itself--namely the Event--and will thus modify the head noun:

(143)



The semantics of the N' will be:

(144) v(<e,x>,N) <<=>> Event(e) & [with(e, a machete) & [destroy(x, the jungle fort, e)]

So the N' destruction of the jungle fort with a machete refers to an Event e where e is an event of x's destroying the jungle fort and e is with a machete.

Note that this prediction is counter to Randall (1984), who claims that these constructions are ungrammatical. This ungrammaticality, she wants to

claim, follows from the Thematic Inheritance Principle insofar as derived nominals, such as destruction may also refer to results. So destruction may, at least marginally, refer to the result of the destruction (the rubble, or in the case of a jungle fort, the hacked-up banana shoots.) This particular interpretation of destruction, which I shall argue in a subsequent section is a different construction from the derived nominal that I have been discussing here, picks up the internal theta role of the verb destroy and as such, it cannot assign it. Given this, and given the Thematic Inheritance Principle, repeated below, we would derive the consequence that, if the two destructions are one and the same, that thematic roles lower on the hierarchy than the Theme, such as the Instrument or the Location should be blocked.

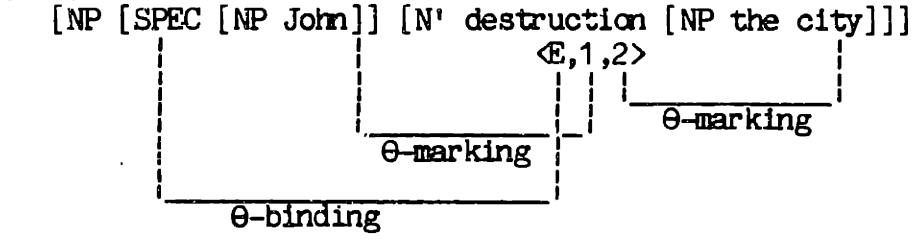
(145) Thematic Inheritance Principle (Randall, 1984; = (56) above)
A category-changing operation which blocks the assignment of a θ-role blocks the assignment of all θ-roles lower on the θ-hierarchy.

However, I have never found anyone who agrees with Randall's judgment that these NPs are ungrammatical, so I suspect that one really can inherit the ability to take an Instrumental or a Location from the underlying verb in a derived nominal. This fact is predicted by the treatment of derived nominals given here.

What discharges the Event position inherited by the nominal? I assume that this is discharged, as is normal in NPs, via binding from SPEC. This is interesting insofar as not only does SPEC bind the Event position, but

the NP in SPEC, if there is one, is theta marked with the external theta role of the verb inherited into the theta grid of the derived nominal; I shall discuss this in more detail below. We thus have a picture which looks like the following[20]:

(146)



Note that under this schema, the external theta role of the verb (the 1), will be discharged last of all insofar as it will be discharged at the same time as the Event, which vies with the external role for last place in this case. But how are the internal and external theta roles discharged precisely? I turn to this in the next two parts.

2.4.1.2 The Discharge of the Internal Theta Role.

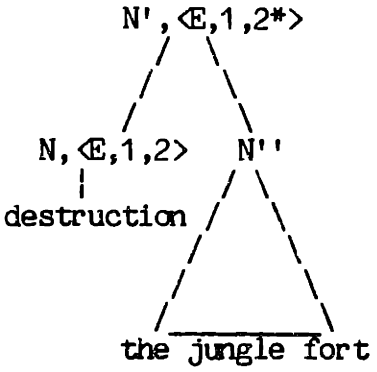
The internal theta role is discharged, I assume, straightforwardly via theta marking, just as a verb discharges its theta role to its object. This is allowed for the derived nominal since it governs its object. The

---

20. I do not mean to imply by this diagram that SPEC and the phrase which it contains are necessarily distinct. The distinction is made in the diagram for expository reasons only.

structure for the N' will thus be:

(147)



Why the internal theta role? The external theta role, as we have assumed, must be assigned last; therefore the only theta role which could be assigned internal to the N' is the internal theta role if there is one.[21]

One major difference between verbs and nominals is that verbs assign structural Case regardless of theta government whereas nouns assign inherent Case under theta government which must be realized in some fashion such as genitive marking or the insertion of a dummy Case-marking preposition. Following Chomsky (1984) I assume that the preposition _of_ may

---

21. It might be objected that with unergative intransitives like _go_ the external theta role can be assigned before the Event is discharged, i.e. internal to the N': 'The going of John.' However, this does not seem to be generally true: *_The acting of John_ ('John acts), *_The walking of John_ ('John walks'), *_The leaving of John_ ('John leaves'.) In fact, as suggested to me by Jim Higginbotham, _the going of John_ may well be an unaccusative construction after all, not an unergative. Note that it cannot refer to an event of John's willful going: *_the intentional going of John_ versus *_John's intentional going_.

be inserted, presumably at S-structure, to realize the inherent Case assigned under theta government at D-structure. That is, whenever, the theta role of a verb is not associated with a preposition (such as to in the case of the Goal in a verb like give), the Case on the theta marked NP in the derived nominal can be realized with of[22]. As the following examples show, an internal theta role receiving NP of a derived nominal is marked with of wherever the corresponding verb phrasal NP is Case marked directly by the verb:

(148)
a. Eric destroyed the jungle fort.
b. Eric's destruction of the jungle fort.

(149)
a. John evaded the CIA for three weeks.
b. John's evasion of the CIA for three weeks

(150)
a. Kyle drank the coffee before I could get to it.
b. Kyle's drinking of the coffee before I could get to it

(151)
a. The USSR invaded Afghanistan.
b. The USSR's invasion of Afghanistan

(152)
a. Fred donated himself to the cannibals' meat fund.
b. Fred's donation of himself to the cannibals' meat fund

(153)
a. Mark expanded into a gas giant.
b. Mark's expansion into a gas giant

---

22. Of course, there are well known cases, such as The general commanded the troops vs. the general's command to the troops where of insertion is not possible. These we discuss below in section 2.4.2.

(154)
a. Edna participated in the pogo-stick race.
b. Edna's participation <u>in</u> the pogo-stick race

Note also that Case assigning properties of nouns are unrelated to that of the verbs to which from which they are derived: Case-assigment is a syntactic property having to do with the lexical class to which a word belongs and having nothing to do with its derivational history; in this way it is utterly unlike theta role assignment, at least as I have suggested. A consequence of this is that there is absolutely no reason why the internal argument of a derived nominal should not appear as the object of the nominal, even when the corresponding verb is unaccusative, hence unable to assign structural Case. This is confirmed, of course, by fairly well-known examples such as the following:

(155)
a. *{It / There} expanded Mark into a gas giant.
b. The expansion of Mark into a gas giant was annoying to the other chess players.

(156)
a. {*It / ?There} arrived John.
b. The arrival of John caused a commotion.

(157)
a. {* It / * There} exploded the bomb.
b. The explosion of the bomb was loud.

So, the internal arguments of verbs may be satisfied in derived nominals by theta marking analogously to the way in which they are satisfied within

VPs. The only difference has to do with Case marking, a difference which is associated with the category difference between verbs and nouns. Before we go on to the somewhat more tricky question of the discharge of the external theta role, let us review a few more properties of the derived nominal which make it different from the verb from which it is derived.

One property is the lack of dative shift in derived nominals, noted and discussed in Anderson (1977, 1979):

(158)
a. Mary gave a car to Harry.
b. Mary gave Harry a car.

(159)
a. Mary's gift of a car to Harry.
b. *Mary's gift of Harry of a car.

(160)
a. John purchased facts for the CIA.
b. John purchased the CIA facts.

(161)
a. John's purchase of facts for the CIA
b. *John's purchase of the CIA of facts

Why should this be? Anderson suggests that the reason has to do with the meaning of the nominals derived from dative verbs. According to Anderson, the "dative noun and its complement name the same entity. Dative nouns tend to name things rather than actions." Some examples, from Anderson, are:

(162) payment, award, delivery, loan, shipment, preparation

Accordingly, Anderson suggests that the reason that dative shift is

impossible in NPs is because in these cases the of is lexically inserted; thus gift of a car is something like 'gift which is a car.' (In our terms, gift is not gift+NOM.) Thus (Anderson, 1977):

> If it can be maintained that all dative nouns have this property [namely that of is lexically inserted, RS] it will be possible to explain why Dative Movement does not apply in NPs. Those nouns with prepositional phrase complements will have no object position to move an indirect object into. From this analysis....it follows that nouns from this semantic class will not allow NP Preposing.

There is, however, strong evidence that the lack of Dative Movement in derived nominals has nothing whatever to do with the referent of the derived nouns of Dative verbs. Dative shift also fails in the following cases:

(163)
a. John is a giver of books to charity.
b. *John is a giver of charity of books.

(164)
a. The giving of books to charity is charitable.
b. *The giving of charity of books is charitable.

(165)
a. John is a purchaser of facts for the CIA.
b. *John is a purchaser of the CIA of facts.

(166)
a. The purchasing of facts for the CIA is immoral.
b. *The purchase of the CIA of facts is immoral.

So, neither giver nor giving can be said to refer to the result of the the act of giving since the former refers to the Actor and the latter to the event or action of giving. Nevertheless, Dative Shift fails.

There have been a number of analyses besides Anderson's which attempt to handle these data; among them see Kayne (1983) and also Rappaport (1983) for a criticism of an earlier version of Kayne's paper plus a proposal within the LFG framework for handling the facts. Mark Baker has suggested to me that the lack of Dative Shift in NP may be related to the fact that nouns assign inherent Case only under theta government, invoking the Uniformity Condition, given above in (24) and repeated here:

(167) Uniformity Condition, Chomsky (1984)

> If a is an inherent Case-marker, then a Case-marks NP if and only if a theta-marks the chain headed by NP.

The idea is that in the case of the Goal or Benefactive argument of a verb such as give, the verb does not theta mark the NP argument directly, but rather does so to the PP, the head of which (a preposition to or for) finally discharges the appropriate role to the NP. Assuming that this property of indirect theta marking carries over to the derived nominal, it will follow from the Uniformity Condition that Dative Shift is not possible since the nominal does not theta mark (directly) the (chain headed by) NP. Hence the shifted NP could not receive Case from the nominal. In the case of verbs, however, nothing blocks the application of a Dative Shift since the shifted NP can get Case from the verb, verbs being structural rather than inherent Case assigners.

Another property of derived nominals which follows from the Uniformity

Condition is the lack of Exceptional Case Marking, as pointed out by Chomsky (1984) (and as also discussed in Anderson (1979), who characterized it as a restriction on Raising.) In a structure such as:

(168) $[_{N'}\ N\ [_{S'}\ NP\ INFL\ VP]]$
$[-fin]$

INFL cannot assign Case to the subject of S' since it is non-finite (or lacks AGR). Case would thus have to come from outside, but this would be impossible given that the head of N' may subcategorize for and hence theta mark the entire S', but that it certainly does not theta mark the subject of S'. Thus the following derived nominals (the primed examples) are predicted bad:

(169)
a. Eric permitted [John to be an aardvark]
a.' *Eric's permission of [John to be an aardvark]
b. Bill believes [Fred to be a coward]
b.' *Bill's belief of [Fred to be a coward]
c. The graduate students considered [the selection committee to be taking too long to decide]
c.' *the graduate students' consideration of [the selection committee to be taking too long to decide]

Presumably also derivable from the Uniformity Condition is the lack of Raising in NPs. This is entirely analogous to the previous cases with Exceptional Case Marking. So, the nominal cannot assign the Possessive Case to its subject if that subject is not theta marked by the nominal. If the NP originated as a downstairs subject, then it will not be so theta marked. The following are thus predicted bad:

(170)

a. *Bill's$_i$ belief [t_i to be intelligent]

b. *Fred's$_i$ consideration [t_i to be a fool]

c. *Eric's$_i$ permission [t_i to be an aardvark]

Of course, theta marked NPs will be more or less freely able to prepose in NP[23]:

(171)

a. The jungle fort$_i$'s destruction t_i (by John)

b. The play$_i$'s performance (by the company)

c. Afghanistan$_i$'s invasion (by the Soviet Union)

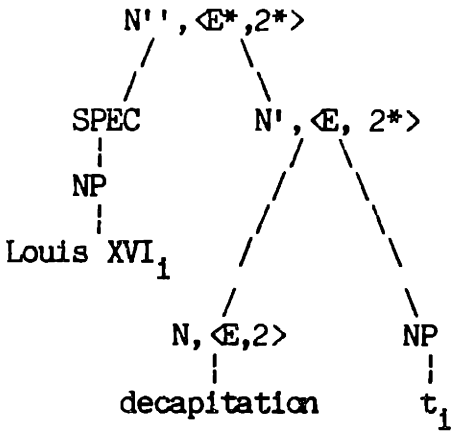
d. Louis XVI$_i$'s decapitation (by the revolutionaries)

The structure of one of these would be given more fully as:

---

23. I say more or less because there seem to be constraints on what kinds of objects may be preposed having to do with "affectedness," as Anderson (1977, 1979) terms it. So in knowledge of algebra the NP algebra is not affected by the action of the verb and it turns out that the NP cannot be preposed: *algebra's knowledge. The precise nature of this constraint would be interesting to work out. Note too that it might be that the constraint is on the kind of theta-role that possessive 's can Case-mark (i.e., be associated with), hence not a constraint on preposing per se.

(172)



So, the NP Louis XVI is generated in the object position of decapitation, where it also gets assigned the internal theta role. It then moves to the subject position of the NP where it ultimately gets assigned Possessive Case. The latter Case assignment is allowed by the Uniformity Condition since the nominal decapitation theta marks the chain, via the trace in the object position.

Of course, there is another assumption which is necessary in order to make all of this work out, and that is that the subject of the NP can be a theta-bar position. That is, if a theta role were to be assigned to the position into which the NP Louis XVI is to be moved, then the movement would be a violation of the theta criterion; it would be ruled out for precisely the same reason as the following structure would be ruled out:

(173) [John$_i$ [saw t_i]]

So (175) cannot mean the same thing as John saw himself since in the

structure above, the chain (John, t) is receiving two theta roles.[24] The same would hold of the chain (Louis XVI, t) if the external theta role of decapitation were to be assigned to the subject position. In fact the external theta role does not seem to be assigned at all; there is no by-phrase (in the version which does not contain the PP by the revolutionaries) and there is furthermore no other place to which we could assume that the theta role is discharged. In particular, we cannot assume, following Baker, Johnson and Roberts (in progress) or Jaeggli (1984) that the external theta role is actually picked up by a passive morpheme as may be the case in syntactic passive, since there is no passive morpheme in derived nominals. What we seem to have to say is that the external theta role is simply not inherited by the derived nominal in this case, a suggestion which is indicated in the above example by the absence of the 1 in the theta grid. Of course, as claimed in section 2.3.2. it is possible for the external theta role to go undischarged and it is therefore possible for the external theta role not to be inherited by a derived nominal. If it is not inherited, then obviously it cannot be assigned to the subject position of the NP either and the subject position can be a theta-bar position. This set of assumptions, apparently necessary in order to be able to account for the possibility of NP preposing, has some interesting consequences to which we now turn.

---

24. On the concept of chains see Barss (1984); Chomsky (1984).

2.4.1.3 On the Discharge of the External Theta Role.

In derived nominals such as the following, where the External Theta Role clearly is assigned, we can assume that the discharge of the external theta role will take place under government from the nominal head:

(174)
a. Mark's expansion into a gas giant
b. Eric's destruction of the jungle fort
c. Mike's evasion of the CIA
d. the Revolutionaries' decapitation of Louis XVI

So, in (174b), for example, destruction governs Eric, and can therefore theta mark it. Furthermore, the External Theta Role will be dischargeable to this position since the condition that the External Role be discharged last is met. Finally, by the Uniformity Condition, since destruction will theta govern the subject, it will also be able to assign inherent Case to it. Thus Eric will show up with possessive.

However, as we saw in the discussion of the NP Preposing examples in section 2.4.1.2. we have to allow that the External Theta Role of the verb remain undischarged, a possibility which we have assumed is true in the general case anyway unless other conditions force its assignment. What then of the following examples?

(175)
a. the destruction of the city
b. the evasion of the CIA
c. the performance of the play
d. the singing of the aria

Is the External Theta Role assigned here? General considerations lead us to suppose that it need not be, yet Chomsky (1984) has argued that in these examples the External Role is assigned, and the assignment takes place to a "hidden pronominal" in the position [SPEC, NP]. Let us call this pronominal PRON and assume that, for some reason, it can only occur in the subject position of the nominal. This assumption seems justified given that the following NPs are ungrammatical with the given interpretation (this fact has been noted by, among others, Roeper (1984), Lebeaux (1984)):

(176)
a. *John's destruction (= John destroyed something or other)
b. *The workers' construction (= the workers constructed something or other)
c. *The policemen's dispersal (= the policemen dispersed someone or other)

Given that PRON could occur in object position of N', there is no reason why the structure of (176a) could not be:

(177) [John [destruction PRON]]

This would give the reading which this NP cannot have.

Getting back to the question of the external theta role, there is no reason why there could not be an empty pronominal in the position [SPEC, NP] which picks up that role, and which acts as a subject of the NP. That this is certainly possible is indicated by the following examples:

(178)
a. The postdocs and the graduate students$_i$ agonized over [the

decision about them$_i$].

b. The FBI and the CIA$_i$ expected [the evasion of them$_i$]

c. Afghanistan and Grenada$_i$ should have anticipated [the invasion of them$_i$]

The interpretation of all of these, according to naive informants, is that the agent of the derived nominal is someone other than the object party (i.e., in the case of (178a), someone other than the postdocs and the graduate students). Assuming that this external theta role is assigned to PRON, which is then chosen arbitrarily to refer to some other person, the pronoun in object position can then be marked as disjoint in reference from the subject of the NP, and will thus have a BT-compatible indexing (see Section 2.2.2., above). It will of course be free to corefer with the subject of the matrix sentence, as indicated in the above example.

The problem comes with examples such as the following:

(179)

a. The FBI and the CIA$_i$ expected [the evasion of each other$_i$]

b. The postdocs and the graduate students$_i$ agonized over [the decision about each other$_i$]

According to some of my informants, and I am inclined to agree, these sentences may have the interpretation that some other person is the evader or the decider, respectively. So (179a) can mean that the FBI expected that somebody or other would evade the CIA and the CIA expected that somebody or other would evade the FBI. Suppose then, that a PRON is

present in these examples. If it is chosen to have arbitrary reference, as the primary interpretation of these examples dictates, then it will follow that there is no BT-compatible indexing for the reciprocal which will yield the interpretation that each other is coreferential with the subject of the matrix clause; this is the case since the nominal head would be a governor for the reciprocal and PRON would act as the subject. The reciprocal would thus have to be locally bound within the NP, which would require, that it be bound to the subject of the NP, yielding the interpretation that, in (179a), the FBI and the CIA expected that some group of people would evade some other group of people, and vice versa. Yet this is not the interpretation. The only structure which would be consistent with the interpretation given to these sentences, plus the fact that they are well-formed with the indexing given, is where the External Theta Role is not discharged at all, where there is therefore no PRON in the subject position, hence no subject, and where each other is therefore free to be bound within the whole sentence rather than only within the NP. The agent of destroying, insofar as it is not syntactically realized, would presumably be chosen to be arbitrary by the pragmatics.

Note that not all speakers agree with these judgments, and neither do speakers who agree on these allow similar constructions in all cases. Thus, one of my informants felt that the following example required that the agent (the invader) be the same as the subject of the sentence:

(180) Afghanistan and Grenada$_i$ should have anticipated [the invasion of each other$_i$].[25]

This might well mean that a PRON was present, coindexed with the reciprocal, and "controlled" in some sense by the subject of the matrix clause.

However, there is a marked contrast between these cases where the subject of the NP, if present at all, is non-overt, and the following cases where the subject is obviously present:

(181)

a. *The postdocs and the graduate students$_i$ agonized over [the search committees' decision about each other$_i$.]

b. *The FBI and the CIA$_i$ expected [the Russian and the East German agents' evasion of each other$_i$]

c. *Afghanistan and Grenada$_i$ should have anticipated [the Russians and the Americans' invasion of each other$_i$.]

In all of these cases the subject of the NP clearly blocks coreference of the reciprocal with the subject of the matrix clause, as predicted by the Binding Theory. There is no disagreement on any of these examples.

---

25. The interpretation my informant gave to this is actually rather surprising, namely that Afghanistan should have anticipated that Grenada would invade Afghanistan and Grenada should have anticipated that Afghanistan would invade Grenada.

This point is, I feel, striking. It simply does not generally seem to be the case that a subject of a derived nominal is required any more than the subject of a picture noun (see also Chomsky, 1984):

(182)

a. Lucia and Donaldo$_i$ espied [nude photos of each other$_i$]

b. Asimov and Sagan$_i$ saw [documentary films about each other$_i$]

c. Edmund and George$_i$ heard [nasty stories about each other$_i$]

If this is a correct assessment, it is certainly as demanded by the theory given that the External Theta Role need not be discharged. It is also demanded by the fact that NP preposing may take place in NP, as pointed out in 2.4.1.2. Of course, there are many interesting questions to be examined. What, for instance, determines, for those speakers who get those judgments, that (179a) is okay without a subject whereas (180) might well require a subject? It might be that pragmatic considerations play a large part: if postdocs and graduate students are agonizing over a decision about them, (themselves, each other) it is very likely, in the current state of the world, that somebody else is doing the deciding. Similarly, the normal expectation (possibly factually false) is that the FBI and the CIA will not be evading each other, but rather that they will be evaded by someone else. Why, if Afghanistan and Grenada are anticipating the invasion of

each other, they should be anticipating that they would be the invaders is less than clear (especially since real world knowledge would seem to force a different interpretation). Nevertheless, whatever determines this may well be extralinguistic. I conclude, then, that the External Theta Role need not be assigned in derived nominals, and hence that PRON need not be present.

Counter to this suggestion, Chomsky has made the following argument to show that a pronominal must in general exist in the subject position of a derived nominal (Chomsky, 1984, p. 244):

> Compare the examples (234):
>
> (234) (i) we thought that [$_S$[$_{NP}$ any attempt to hurt each other] would fail]
> (ii) we thought that [$_S$[$_{NP}$ any attempt to hurt us] would fail]
>
> In (i), it is <u>our</u> attempt that we think will fail, while in (ii) it is someone else's attempt, consistent with the requirements of the Binding Theory if we assume that the Determiner of the subject off the embedded clause contains a hidden pronominal.

But these examples seem dubious simply because, on independent grounds we can derive the presence of the hidden pronominal in cases where control is involved. That is, a nominal such as <u>attempt</u>, when it takes an infinitival complement, always acts as a domain for control in that its subject is consistently construed as being identical to the subject of the embedded clause. But this just means that PRON must be present in case there is an

infinitival as a complement to such a nominal. These examples certainly do not argue that such a pronominal must generally be present in derived nominals.

2.4.1.4 Some Problems with and Predictions of the Model.

In this section I discuss some random facts which bear on the model for derived nominals and theta role assigment presented here.

One possibly problematic fact for the whole analysis presented here is the fact that by phrases seem to be able to discharge the External Theta Role of the nominal:

(183)
a. the destruction of the jungle fort by John
b. the evasion of the FBI by the hoods
c. the baking of the cake by the chef

This would be problematic if the by-phrase discharges the external theta role and if it is within N', since it would seem that the External Theta Role would be getting discharged before the Event position, which is bound by the SPEC. We could perhaps allow this to happen by relaxing the definition of External Theta Role to allow that the Event does not count; in fact, it does not seem that the precise manner or location of the

Event's discharge is really crucial for the External Theta Role.[26] What is crucial is the relationship between the External Theta Role and the Internal Theta Roles, namely that the former must be discharged after the latter, a relationship which we can assume is preserved in the examples in (183).

It is also conceivable, however, that the by-phrase does not hang off of N'. In this case, the External Theta Role would be discharged at the same time as the Event, namely under N''. This supposition is certainly consistent with the following data:

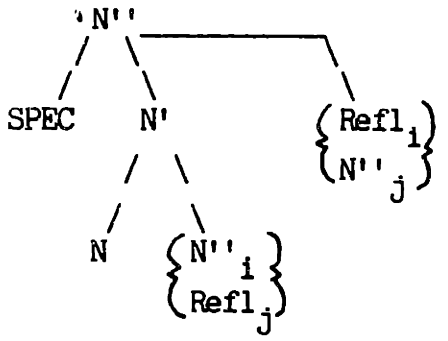
(184)
a. the destruction of himself$_i$ by John$_i$
b. ?*the destruction of John$_i$ by himself$_i$

Allowing that the PP dominating the by-phrase does not count in the determination of c-command (and is thus like a to goal phrase, see Reinhart (1976)), the structure of the NPs would be:

---

26. We did assume (section 2.3.2.) in -er nominalization that the undischarged Event could block the discharge of the External Theta Role under identification, in the intransitives, to the theta role of -er. Nevertheless, this turns out not to be crucial since even if this were to happen, the interpretation of the forms would be the same.

(185)



(i indexing *, j indexing okay)

This structural asymmetry would help explain the observed binding data, assuming Reinhart's (1976) definition of c-command: (184a) is okay since himself is bound by a c-commanding anaphor. (184b) is bad since himself is unbound, as John does not c-command it. However, if the by-phrase were instead to hang under N', we could not explain the asymmetry.[27] In any event, I expect that the problem with by-phrases can be cleared up pursuing one of these two lines of investigation.

One interesting question arises when we consider what happens if we base generate PRON in the object position of NP. We have claimed that it cannot remain there, but there is no reason to assume it cannot be generated there; presumably its absence from the object position may, like PRO, be determined by S-structure conditions. Therefore, provided that the subject

---

27. See Giorgi (1984), for a similar argument from Italian leading to the same conclusions for that language.

position of the NP is a theta bar position--i.e., the external theta role is not assigned--we will be able to move PRON from the object position to the subject position yielding a structure like the following:

(186) [$PRON_i$ [N t_i]]

This predicts, correctly, that the following event nominals with no overt arguments should be fine:

(187)

a. Destruction with a machete is fun.
b. Construction with pile drivers can be hard on the nerves.
c. The performance lasted three hours.
d. Dispersal with tear gas is the preferred method.
e. Confirmation usually takes about three weeks.

(Some of these, such as (187a,b) sound better for some reason without the definite article.) Note that this analysis of event nominals lacking overt arguments necessarily links these constructions to cases of NP preposing. We might therefore expect that examples where NP preposing is impossible, such as the cases cited by Anderson where the object of the action is not affected by the action, should be ill-formed with these constructions[28]. This is correct, I think, although the judgments are delicate:

---

28. Though this is not so clearly correct insofar as we might imagine, as suggested previosly in footnote 23, that Anderson's constraint should be construed as a constraint on Case-assigment by 's rather than one on NP movement. In this Case, since PRON presumably does not need Case (since it has zero phonological content), we would expect that no restrictions would hold of the types of nominals where PRON may originate in the object position.

(188)
a. *the cat's pursuit by Sally
a'. ?Pursuit with a machine gun is often not recommended.
a''. Pursuit of pussy cats with machine guns is not ethical.

b. *Mary's discussion by John
b.' ?Discussion over coffee is often nice.
b.'' Discussion of Mary over coffee is the only way to resolve this problem.

Summarizing so far, we have argued that the discharge of Internal Theta Roles proceeds within NPs as it does in verbal projections, and that the discharge of the External Theta Role is optional. In this latest section we have discussed a couple of minor problems with and consequences for the analysis. In the next section, I discuss synthetic compounds with derived nominals.

2.4.1.5 On Synthetic Compounds with NOM.

Since NOM is an affix whose only syntactic specification is that it subcategorizes for a verb and makes a noun, there is no reason why it could not attach to a verb which is formed from a noun compounded with a verb; synthetic compounds with derived nominals should thus exist. This expectation appears to be confirmed by the following data:

(189)
a. Jungle-fort-destruction with a machete usually takes about three hours.

b. Termite-annihilation with insecticide is of grave concern to the Society for the Preservation of Aardvark Habitats.

c. Play-performance should not take place in the back alley.

d. Problem-solving by computers is faster than with a slide rule.

e. Dog-detection using advanced radiometric techniques is seldom faster than the old "sausage lure" method.

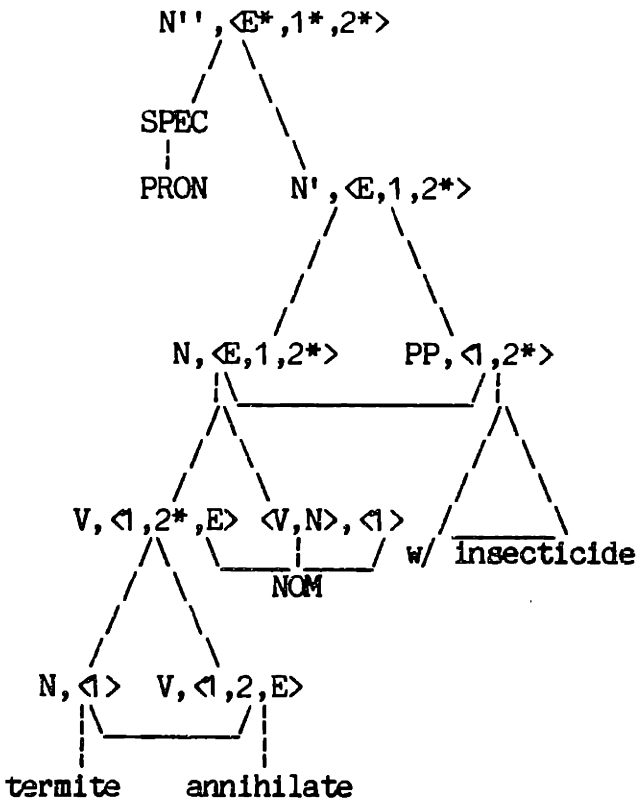
f. Friend-acknowledgment is usually expected of a dissertation writer.

g. Grain-importation on ships is invariably the method chosen.

h. King-decapitation is a method used by revolutionaries who are tired of the current Head of State.

Some of these examples are perhaps more marginal than others but they all seem to be interpretable in the appropriate way, namely that the left member of the compound is picking up the internal theta role of the verb. I therefore propose that the structure for, for example, (189b) is:

(190)



So, as with synthetic compounds in -er the left sister discharges the 2 position of the verb. The 1 position of the affix will discharge to the Event position of the verb, thus meaning that, as in non-compounded derived nominals the Event will become the referent of the nominal, and also predicting, as indicated in the above example, that Instrumentals and similar modifiers of Events should be possible in these constructions. The semantic interpretation of the derived nominal compound will be:

(191) $w(\underline{e},\underline{N})$ <<=>> $Event(\underline{e})$ & $\exists\underline{x}$ [$termite(\underline{y})$ & $annihilate(\underline{x},\underline{y},\underline{e})$]

The internal theta roles, following the Projection Principle, will have to be satisfied within the compound. This is correct as shown by the following examples (some of which have been already discussed in Section 2.3.5.3.):

(192)
a. *child-giving of candy
b. *candy-giving to children
c. *meat-donation to the cannibals
e. *cannibal-donation of meat

Finally, the External Theta Role will be satisfiable outside the compound since it is not restricted by the Projection Principle to occur inside the compound. Thus overt external arguments ought to be allowable. This is correct:

(193)
a. Their city-destruction should always be brought to the attention of the proper authorities.

b. Edna's termite-mound-annihilation with a can of Raid® was disturbing to the aardvarks.

c. Politician-defamation by this reporter is becoming common in this newspaper.

d. The mathematician's problem-solving with a pencil and paper took a long, long time.

Some of these sound perhaps a little strange since, for some reason, it seems preferrable to interpret these compounds as referring to generic rather than specific events; that is. city-destruction seems to be better interpreted as the class of events of destroying of cities, rather than a

particular instance of such destruction. Other than that, the examples seem fine.

These compounds are, of course, Bracketing Paradoxes. The particular spellout of NOM depends upon the verb chosen, and it thus seems appropriate to claim that this affix brackets with the verb alone in the mapping to PF. We thus have the following mismatch between the syntactic and the phonological representations of the following example:

(194)
[[TERMITE ANNIHILATE] NOM]
[termite [annihilate ion]]

I will omit the proof that this is possible since proofs for entirely analogous situations have been given in the last chapter. Note in particular that whatever phonological spell out of NOM is chosen by the verb will show up in the compound. Thus, the event nominal of solve is not solution but rather solving; solution is a result nominal (see the next section). So if we have a structure of the form in (195a) we expect to get a phonological form like that in (195b) not like that in (195c). This prediction is confirmed:

(195)
a. [[PROBLEM SOLVE] NOM]
b. problem-solving
c. *problem-solution

Examples of this form bring up an interesting point. If we assume that the analysis of compounds presented in this section is correct then we are

forced to the conclusion, given these examples, that the affix we are attaching syntactically is an abstract affix such as NOM, rather than a more specific affix such as -tion. The reason for this is as follows. Suppose that we wanted to claim that Event nominalization could be mediated productively at the syntactic level by a large number of specific affixes, including -tion. We might then say that the only syntactic specification on these affixes was the specification that we have assumed is marked on NOM, namely that they subcategorize for verbs and produce nouns. The particular affix would then be checked in the PF-component for well-formedness; in particular -tion would pass the test at PF if it is attached to any of a large class of latinate verbs. Now, consider what happens when we have a verb such as solve. When -tion, as opposed to -ing attaches syntactically to this verb alone--that is, solve uncompounded with anything--we get a special interpretation, namely that of result. Furthermore, this is the only interpretation possible so it will be necessary to list somewhere that the syntactic construction [SOLVE TION] must get the result interpretation; this is of course needed on anybody's theory since this is presumably not a predictable fact. So far so good, but what happens when we attach -tion to a compound with solve such as a construction like [PROBLEM SOLVE]? Since [PROBLEM SOLVE] is not the same as SOLVE there is no reason why any special rules of interpretation should be invoked here and the semantics should be perfectly happy interpreting this construction as meaning the same thing as problem solving. In the

mapping to PF the phonological forms of SOLVE and TION will rebracket with one another and they will pass the PF filter requiring that -tion be attached to a latinate form and we will thus have problem-solution as a perfectly well-formed word referring to the event or action of problem solving. This, of course is completely wrong. Thus, we seem forced to the alternative conclusion, namely that there is a more abstract event nominalization affix, which we have called NOM, and whose phonological shape is only determined in the mapping to PF, the determination being sensitive to the particular verb in all cases. Of course, one could knock down this whole argument if one were to argue, on the contrary, that one cannot assume that the compounds discussed in this section are synthetic compounds. Yet they behave like synthetic compounds, as I have argued, and there seems therefore to be no principled reason to deny that that is precisely what they are. We therefore seem led to the conclusion that NOM is a real syntactic entity, and spells itself out phonologically with the affixes -ing, -tion, -ment, -ance...., each case depending upon the particular verb chosen[29]. This conclusion has non-trivial implications for acquisition, which we turn to in the next section (2.4.2.). Before we turn to that, however, I would like to discuss a couple of other kinds of data,

---

29. Though it generally seems to be true that -ing is available with this reading for most verbs. As has been often noted, however, there seems to be some sort of blocking going on insofar as if, say, -tion is used with a particular verb, then -ing sounds distinctly odd: c.f. ??John's destroying of the city, vs. John's solving of the problem.

discussed in the literature which have a bearing on the analysis presented here.

The first fact is discussed by Fabb (1984, pp. 187-8). He notes the following contrasts.

(196)

a. grain non-importation, *grain non-import{ing / er}

b. industry non-regulation, *industry non-regulat{ing / er}

c. food non-spoilage, *food non-spoil{ing / er}

He uses these facts to argue that there is a difference, contrary to what I have suggested here, between (latinate) derived nominal compounds, and familiar cases of synthetic compounds in _-ing_ and _-er_. So, in the _-ing/-er_ cases, _non-_ would intervene between the verb and the noun which would mean that the noun could not get Case and the construction should thus be ruled out under both Fabb's and my analysis of synthetic compounds. If the examples with _-tion_ and other latinate affixes are at least marginal then this suggests that these might actually not be synthetic compounds after all.

To my ears, at least, these examples are distinctly odd; even Fabb admits that they are rather marginal. What is even more interesting is that the one example which sounds at least plausible, namely _government non-intervention_, has the interpretation that _government_ is picking up the

External Theta Role of the verb, rather than an internal theta role; after all, intervene does not have an internal theta role. Thus, this could not be a synthetic compound anyway. What it could be is a root compound with the left member being assigned the "semantic role" (see Lieber, 1983) of Agent. I believe that this, in fact, is what is going on in all the examples which Fabb provides; that is, the left member forms a root compound with the non-N construction, and is interpreted "semantically" as picking up the internal theta role, which is presumably possible for semantic arguments[30]. This assumption seems to be confirmed by the fact that in general the intervention of non- is not possible, and we can therefore reasonably assume that the interpretation of such forms is always going to be by some marginal roundabout means:

(197)
a. *Aardvark non-annihilation is preferred.
b. *Dog non-detection is a real possibility.
c. *Meat non-consumption is usual.

I therefore do not believe that examples of this form demonstrate that (latinate) derived nominal compounds are necessarily not synthetic compounds.

---

30. As examples like CIA-informant (= 'one who informs the CIA'), in fact, suggest; -ant is an affix which selects for some verbs but not for the majority, and it is therefore likely not to be the case that CIA is syntactically bracketed inside -ant. Therefore, it must be outside, forming a root compound with informant, and thus picking up the internal theta role as a semantic argument.

Let us turn now briefly to some examples discussed in Pesetsky (1985) and Selkirk (1982). Both of these researchers note that examples such as the following are bad (from Pesetsky, p. 236):

(198)
a. *Weather changing is typical of Western Massachusetts.
b. *Time elapsing worries philosophers.
c. *Heart beating is a good sign of life.

In all of these examples the theta role obviously intended for the left member of the compound is the External Theta Role. This assignment seems to be impossible. Selkirk captures this straightforwardly with the following stipulation (p. 35):

(199) The external argument of a lexical item may not be satisfied in compound structure.

Pesetsky has a somewhat different and more complicated way of deriving this fact. I will not discuss his analysis here but rather refer the reader to Pesetsky's paper.

How do we derive this fact? We can assume first of all that the following structure would be impossible for these forms with the desired interpretation:

(200) [[N V] NOM]

If there is an internal role to discharge, then by the Projection Principle that role must be discharged within the compound and therefore the External Argument cannot occur there especially given the fact that it must be

satisfied after all internal arguments. If the verb is monadic as in the above examples, then it will not be a Case assigner anyway, so the N would not be getting Case. The construction would therefore be ruled out by the Case filter. The only other possible structure would be:

(201) [N [V NOM]]

The construction would thus be a root compound rather than a synthetic compound, and the left member would be interpreted as the semantic Agent (following Lieber (1983), again) of the action expressed by the verb. However, we would expect that the interpretation of semantic Agency for these forms would be as sporadic and idiosyncratic as the interpretation of semantic arguments usually is. In fact, there are forms, which Pesetsky notes, where the interpretation seems fine:

(202)
a. Dog barking can be eliminated with a muzzle.
b. Child dancing is forbidden.
c. Student swimming is allowed only on Thursdays.

The situation seems entirely parallel to the situation with lexical passive compounds; some allow the Agent, some don't (see also Lieber (1983)):

(203)
a. Indian-made pottery goes on sale at half-price tomorrow.
b. ??Woman-made clothing is usually good.
c. ?Farmer-killed chickens are nice to eat.
d. *Pianist-played sonatas sometimes sound nice.

Note also that this lack of predictability seems to carry over to nominalizations in (phonological) affixes other than <u>-ing</u>:

(204)

a. Government intervention is sometimes a necessary evil.
b. ?Student destruction is commonplace these days.
c. ??Dog expansion is an inevitable consequence of overfeeding.
d. *Government deletion is the ultimate fate of embarassing evidence.
e. *Industry pollution is a fact of life in the 20th century.

(Be sure always to pronounce these with compound stress (i.e., with the stress on the first member of the compound) since these forms can also be interpreted syntactically as being of the form Adjective Noun, where the lefthand nouns government, student, and so on, are being used as adjectives.)

I suspect, therefore, that the possibility or impossibility of interpreting the left-hand member in the examples discussed above is going to be subject in part to the vagaries of semantic argument interpretation. That is to say that I have explained why these forms cannot be synthetic compounds (hence reliably and predictably interpreted with the left-hand member as the external argument), but I have not explained why in particular cases the interpretation is or is not possible as a root compound. This may seem unsatisfying insofar as Selkirk and Pesetsky both, after a fashion, provide explanations for this fact. On the other hand, since the phenomenon extends beyond the domain of these Event nominals, I suspect that their explanations are false.

2.4.2 Not All Derived Nominals are Created Equal.

The reader who is at all familiar with the field of derived nominals will undoubtedly have been annoyed for some time that I have virtually completely ignored the following well-known fact: It is simply not true that all nominals related to verbs allow for the internal argument of the verb to be expressed with an NP governed by <u>of</u>; it is quite often the case, rather, that a different preposition is required, one which it can be fairly convincingly argued is selected on the basis of its semantic appropriateness for the task of assigning the particular theta role. Some fairly well-known examples (some of these from Rappaport (1983)) are given below:

(205)

a. Herbie promised Louise to write.
a'. *Herbie's promise of Louise to write
a''. Herbie's promise <u>to</u> Louise to write

b. The general commanded the troops to leave.
b'. *The general's command of the troops to leave
b''. The general's command <u>to</u> the troops to leave

c. Seymour permitted Irving to take a break.
c'. *Seymour's permission of Irving to take a break
c''. Seymour's permission <u>to</u> Irving to take a break

d. John solved the problem.
d'. *John's solution of the problem
d''. John's solution to the problem

e. John fled the city.
e'. *John's flight of the city
e''. John's flight from the city

f. The soldiers entered the city
f'. *The soldiers' entry of the city
f''. The soldiers' entry into the city.

Why should these examples exist? What do we need to say about these deverbal nouns that would predict that they behave differently from the examples discussed in the last section? Clearly what we need to say is that these deverbal forms are somehow irregular--that is, not produced by the productive affixation of a nominal affix such as NOM. Then, if they inherit the thematic roles at all from their base verb it will only be sporadically that they do so. But is there evidence that forms such as those discussed above are truly different from the NOM nominals? There is: in fact, the evidence is quite striking.

First of all, not all of the forms given above even refer to events; the ones that do not could not be cases of the affixation of NOM. For those forms which do not refer to events, we would expect that Instrumentals and other modifiers of the Event position would be unavailable in the nominal. This is correct:

(206)
a. The general commanded the troops with gusto.
a'. *The general's command to the troops with gusto impressed his foes.
a''. The general's commanding of the troops with gusto impressed his foes.

b. John solved the problem with a calculator.
b'. *John's solution to the problem with a calculator was unsatisfactory.
b''. John's solving of the problem with a calculator was unsatisfactory.

In fact, forms like solution, promise and command are result nominals[31].

Other examples do seem to be reasonable:

(207)

a. John's flight from the city in a car.....
b. The soldiers' entry into the city with a battering ram...
c. The sailors' attack on John with knives....
(c.f. The sailors attacked John with knives.)

So, flight, entry, attack and presumably others do refer to events. Nevertheless, following our assumption that these are not the product of NOM affixation, we can give the following lexical entries for the forms promise and flight (assuming that promise is a three-place verb with 1=agent, 2=goal, and 3=theme):

(208)

$N, \langle 1_i \rangle$
|
$V, \langle 1, 2, 3_i, E \rangle$
|
promise

$N, \langle 1_i \rangle$
|
$V, \langle 1, 2, E_i \rangle$
|
flee

So, the meaning of the noun promise is got in part by indexing the 3 position of the verb; remember that theta indexing does not discharge any roles, but rather merely marks which role is being picked up by the derived

---

31. This is presumably true, as Anderson (1977, 1979) suggests, of many "Dative Nouns" such as gift, which, nevertheless, take objects with of. However, it seems plausible to say here that of has a semantic function here and actually means something like 'consisting of': thus a gift of $5 is a gift consisting of $5. This is basically what Anderson claims too.

word. Similarly, flight is indexed with the Event position. I am, of course, ignoring as unimportant the affixes -0___ and -t which do, presumably have a syntactic representation.

So, if the arguments of the verb are to be "expressed" at all it will only be by choosing an appropriate preposition to theta mark as well as Case mark the NP. In that way, it will be possible to recover the intended meaning of the NP. So, in the general's command to the troops, to marks the fact that the internal argument of the verb is, in fact, a Goal. From is picked for flight for similar reasons. This is similar to Rappaport's claims about the function of prepositions in derived nominals in general. The external theta role of the verb does seem to be readily identifiable with the possessive, but this is presumably because the possessive takes a rather wide range of semantic values in any case (see Williams, 1982, in particular, on this)

If these forms are irregular in that they are not formed by a productive syntactic process such as NOM affixation but rather have their syntactic forms listed, it should follow that synthetic compounds with these forms will be out, for the same reason as forms like *white elephantine, discussed in Chapter 1, are out; the nominal form is related to a particular verb, not just any verbal node which happens to contain this verb. This is in marked contrast to NOM which can apparently affix to any verb. The prediction seems to be borne out:

(209)
a. ??Troop-command is fun.
b. *A friend-promise should not be broken.
c. *Dog-permission is not permitted.
d. *Problem-solution is often hard.
e. ??City-flight takes about three hours.
f. *City-entry should take place through the West Gate.

Note that I am not claiming that a phonologically irregular derived nominal is necessarily irregular--i.e., must be listed--at the syntactic level. An example which may not be irregular in the relevant sense, even though it is irregular phonologically is sale. Evidence for this is, firstly, that it takes objects with of and allows for modifiers of Events:

(210)
a. John sold the apples on the lawn.
b. John's sale of the apples on the lawn

and, secondly, it allows apparently productive compounding which appears to be synthetic:

(211) book-sale, car-sale....

What does appear to be true, however, is that there is evidence that nominalization does not come in one flavor: there is productive and semantically transparent nominalization with NOM, and then examples of nominalizations which are not productively produced and which are rather just listed as being derived from particular verbs. Of course, the affixes may well be phonologically the same in either case. For instance, -tion is used both productively, as in destruction, construction (Event reading), and so on, and unproductively, as in solution construction (Result

reading), etcetera. This, given the viability of the entire analysis presented here, is all the more evidence for the autonomy of syntax with respect to phonology and of phonology with respect to syntax, even at the level of "word formation."

But there is a question here, lurking in the background. How is such knowledge acquired? How do I tell, given that there is no reliable phonological indication of the fact, that it is the productive NOM which is involved in a particular word, or that the word is semantically idiosyncratic and must simply be listed? I believe the answer must be that the learner assumes that the affixation is productive unless there is overt evidence to the contrary. That is, given that the learner can ascertain the fact that a given noun is related to some verb--a non-trivial problem, but a separate one from the problem which we are considering, he will continue to assume that it is somehow predictable in its meaning unless he has overt evidence that he must simply list it. Given that NOM exists as the productive affix referring to events, he might well just assume that this is the reading associated with the noun. Overt evidence to the contrary could consist in any number of things such as the use of semantically contentful prepositions to express arguments of the nominal, or the obvious intended referent of a particular noun as being the result of an action.

How would the learner discover the existence of NOM in the first place?

Certainly derived nominals referring to events are not uncommon cross-linguistically, so it may be that the existence of such nominals will be expected given UG. If this assumption seems unappealing, then one can always assume that sufficient overt evidence for the existence of a "process" of event nominalization--evidence such as exposure to examples like John's destruction of the city--is enough to key the learner in to the existence of an affix with the lexical entry we have assumed for NOM.

Needless to say, this is all speculation. I know of no evidence from acquisition which would bear on this question. In any event, it would seem to be an area worthy of further investigation.

2.4.3 A Discussion of Lebeaux (1984).

Lebeaux in a recent paper presented an analysis of derived nominals in English which differs in a number of ways from the analysis presented here. Assuming Pesetsky's (1985) theory of QR in morphology, he suggests that when a derived nominal, such as destruction looks as if it is assigning the thematic roles of the verb, it really is the verb that is assigning the thematic structure, the affix having LF moved to a position adjoining the appropriate phrasal projection of V. Consider the following examples:

(212)
a. the destruction (extended over several blocks)
b. the destruction of the city (was terrible)
c. John's destruction of the city (was ruthless)

Lebeaux assigns the following LF representations to these, respectively:

(213)

a. [the [[destroy$_V$ tion]$_N$]$_{N'}$]$_{N''}$

b. [the [[destroy$_V$ [the city]]$_{V'}$ tion]$_{N'}$]$_{N''}$

c. [[John's [destroy$_V$ [the city]]$_{V'}$]$_{V''}$ tion]$_{N''}$

Lebeaux is able to derive a number of facts from this assumption. First of all, as suggested by the above examples, and as also noted in my discussion of the topic in the preceding sections, the subject may only be assigned a theta role if the object has its theta role assigned, and this follows from the analysis insofar as the affix may raise to adjoin to V'', where the subject theta role may be assigned, only if it has also "skipped over" the V' where, by the Projection Principle, the internal object is required, and thus gets its theta role. Under my analysis, as we have seen, this set of facts falls out from the definition of the External Theta Role; it is always the last to be discharged so that the internal role is guaranteed to be discharged before it, and in fact it is free to not even be inherited by the derived nominal in which case it is not there to be assigned by the nominal in the first place.

Another fact which Lebeaux's analysis accounts for is the fact that the cases where one can clearly tell that verbal theta roles are being assigned have process readings, whereas the result nominals-- those where no roles are assigned, have no process reading. This follows if we assume, as

diagrammed above, that the verbal theta roles are assigned just in case there is a verbal projection present. If there is a verbal projection, then we can safely assume that a processual reading is present and there ought to be overt evidence such as the existence in these forms--but not in the result forms--of temporal adjuncts. This is correct:

(214)

a) The presentation of this material while groggy from cold medicine would be ill-advised.

b) *The presentation (= result) while groggy from cold medicine was boring.

But these examples are compatible with the analysis I have presented here insofar as temporal adjuncts are plausibly modifiers of events. That is to say, if John presented the material while groggy from cold medicine then there was an event e where e was a presenting of the material by John and e took place while (John) was groggy with cold medicine. But if this is right, then this falls out from my analysis since, as we have seen, nominals which assign verbal theta roles also refer to events, which thus predicts that modifiers of events ought to be allowable. However, result nominals clearly do not refer to events, predicting that event modification would (obviously) be impossible.

We have seen, then, that Lebeaux's major predictions really fall out just as easily under my analysis. Given that I have argued against Pesetsky's morphological QR on other grounds (see chapter 1) it seems plausible to suggest that QR is not what's involved here either, since

there seems to be a perfectly reasonable analysis of these facts which doesn't involve it.

One might wonder why Lebeaux's analysis would not predict that verb-phrasal phenomena such as dative shift and Exceptional Case Marking would occur in derived nominals. I suspect that this is not really a problem for him, however, insofar as these processes are associated with verbal projections in the syntax and not at LF. If we can assume that dative shift has something to do with Case assignment, as suggested above (p. 190-191), then both of these phenomena are presumably related to the particular Case-assigning properties of verbs; verbs, being structural Case assigners, are not constrained by the Uniformity Principle. Derived nominals, for Lebeaux, are only verbs at LF, where issues of Case assignment are presumably irrelevant.

2.5 Concerning Self-analysis.

In previous sections of this chapter I have argued that obviously syntactic conditions such as Theta-role assignment and Case theory are relevant to the determination of well-formedness in structures which are often considered to be morphological. In this, I agree in a large part with the general thrust of Fabb (1984). In this section I would like to consider some data which suggest that the Binding Theory applies in

morphological constructions—apparently the same way as it applies in syntax. The data I shall be using as evidence are "compounds" formed with the English prefix self-.[32]

The data we will be discussing in this section are exemplified in the following:

(215)
a. John is a self-admirer.
b. Self-destruction in this situation would be advisable.
c. John is self-destructive.

I am going to be making the following assumptions: first of all, I will assume that self is, like its syntactic counterparts (i.e., himself, herself...) an anaphor and therefore is subject in the general case to condition A of the binding theory. I shall also assume the Binding Theory applies to these constructions in precisely the same way that it applies in the syntax. In particular, I shall be assuming that self- in such forms is in general syntactically bound, and not lexically bound as it might appear. The benefit of this move will be that it is perfectly possible to make interesting predictions about the way that these constructions behave without any special assumptions about the Binding Theory in lexical constructions. Finally I shall be assuming that self, being a prefix rather than a noun, does not require Case. The importance of this

---

32. Many of the data in this section were investigated in collaboration with Kyle Johnson. Thanks to Kyle for much useful discussion.

assumption will be clear when I discuss adjectival forms like self-destructive, and also cases where self has apparently attached to a noun.

For now, observe that there is at least phonological evidence that self really is a prefix. Self "compounds" do not receive compound stress, so that the stress falls consistently on the second member.

(216)

1 1 1 1
self-destruction, self-evidential, self-admirer, self-explanatory

In the above, a 1 marks the main stress in the "compound". Note the contrast with clear compound cases which usually have the stress on the first member:

(217)

1 1 1
city-destruction, cat-admirer, dog-house,....

There is, of course, a noun self, but this is presumably a different lexical item. In fact, a form such as self-denying is ambiguous if one does not know the stress; a self-denying philosopher is a philosopher who denies himself (the pleasures of life, etc.) if said with stress on denying whereas he is a philosopher who denies the existence of the entity termed the Self if said with stress on the first member of the compound.

It also appears that self can be sister to a member of any category, and that the resulting form is of the same category as the form to which the

affix attaches. I shall argue that the following three cases are examples of attachment of self to a verb, a noun, and an adjective respectively:

(218) self-admirer, self-love, self-destructive

This means that self is of the type $\langle 0,0\rangle$, that is it attaches to anything and does not percolate a category specification. We therefore give the lexical entry for self as follows:

(219)

$$\text{SELF}' = \langle \text{SELF}_{\langle 0,0\rangle}\langle 1\rangle, \text{self}\rangle$$

pre(self)
Anaphor(SELF)

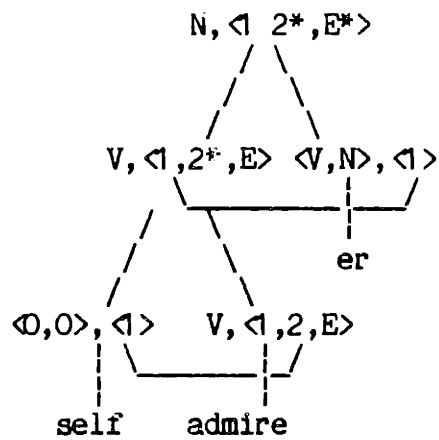
(I assume a predicate **Anaphor** which is true of **x** if and only if **x** is an anaphor.) Note, too, that since self is here given with its own theta-role which it identifies with the verb's internal theta role, it is formally a predicate. This is not unnatural when we note that it will eventualy pick up its semantic content from its antecedent and therefore be a predicate with precisely the same content as the antecedent.

Let us turn now to a discussion of the various forms introduced above and see how the lexical entry for self given above, along with assumptions of the Binding theory and Case theory serve to make the correct predictions.

2.5.1 -Er forms with self.

Following the discussion in previous sections, I assume that the structure for a form such as self-admirer is as follows:

(220)

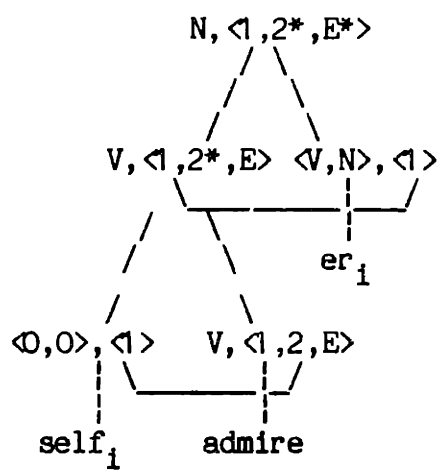


So, theta roles are assigned under identification between the internal theta role position of the verb's grid and the 1 position of the affix's grid. According to the explicit assumption made above, Case will not be assigned here given that self does not require Case, but that is actually irrelevant for this example.

What about the anaphoric status of self? According to the Binding Theory (from Chomsky (1984), and given above in (20-1), self will have to have a BT-compatible indexing within its minimal Complete Functional Complex (CFC). What is the minimal CFC of self? Since it is being assigned an internal argument of the verb admire, the minimal CFC of self

will be that domain in which all of the arguments of the verb are satisfied, including the subject. Since the External Theta Role is finally discharged by SPEC in the NP, the CFC for these examples will be the entire NP. What is the antecedent of self? I suggest that the antecedent is -er; the suffix has all of the right properties in the sense that it both c-commands the anaphor, and it identifies its theta role with the external theta role of the verb, thus giving us the right reading. The structure will thus be as follows, filling in the appropriate indexing:

(221)



Let us assume that this kind of coindexation of antecedents with anaphors such as self introduces an equation of the form x=y into the semantic formula, where x is the antecedent and y is the anaphor. This equation just identifies the two points in the phrase marker as denoting the same object. The semantic value of a noun such as self-admirer will therefore be as follows:

(222) $v(\underline{x},\underline{N})$ <<=>> Actor($\underline{x}$) & $E(\underline{e})E(\underline{y})$[admire($\underline{x},\underline{y},\underline{e}$)] & ($\underline{x}=\underline{y}$)

We can simplify this expression by making use of the equation:

(223) $v(\underline{x},\underline{N})$ <<=>> Actor($\underline{x}$) & $E(\underline{e})$[admire($\underline{x},\underline{x},\underline{e}$)]

So a self-admirer is some $\underline{x}$ such that $\underline{x}$ is an Actor and more specifically there is an $\underline{e}$ such that $\underline{e}$ is an event of $\underline{x}$'s admiring $\underline{x}$, which is, of course, the correct interpretation.

One question which might arise is, given that the CFC is really the whole NP, why self would have to be bound by the -er. Why could it not be bound by the subject of the NP, if there is one? In fact, it cannot, as examples of the following form show:

(224)
a. *John's$_i$ self$_i$-admirer is late today.
b. *Frank's$_i$ self$_i$-redeemer is a fool.
c. *Charlie's$_i$ self$_i$-flagellator is enjoying himself.

In fact, all of these are bad with the indicated indexing; John's self-admirer simply cannot mean somebody or other who admires John, but rather refers to a self-admirer who happens to be in some sort of relationship, expressible by the possessive, to John.

Yet there is evidence that for some reason or other the subjects in these cases are not within the CFC of the self prefix; that is, they are somehow outside the theta domain for the anaphor as far as the Binding

Theory is concerned. This can be shown with clearly syntactic reflexives which are receiving the internal theta role of the verb from the derived Agent noun rather than from the verb directly via compounding:

(225)
a. *John's$_i$ admirer of himself$_i$ is late today.
b. *Frank's$_i$ redeemer of himself$_i$ is a fool.
c. *Charlie's$_i$ flagellator of himself$_i$ is enjoying himself.

So, (225a), if it means anything at all, refers again to a person who admires himself and who is somehow related to John. Actually, I find forms such as admirer of himself fairly odd anyway, as do a number of other speakers whom I have asked (though see Giorgi (work in progress) for data from Italian in which such forms are apparently good.) In any event, there seems to be something blocking coreference between an anaphor which is the internal argument of a form such as admirer and the subject position of the NP in these cases. I do not have an explanation for this, though presumably the N' is somehow acting as a domain for the Binding Theory in these constructions, for whatever reason. Similar judgments carry over, I think, to picture NPs also: so, John's story about himself is fine as long as John is the one who is telling the story, but for some reason the judgment for me degrades if someone else has told the story, and John just happens to own a written transcript of the telling. In any event, whatever the explanation for these forms is, it will surely carry over to examples of the form John's self-admirer.

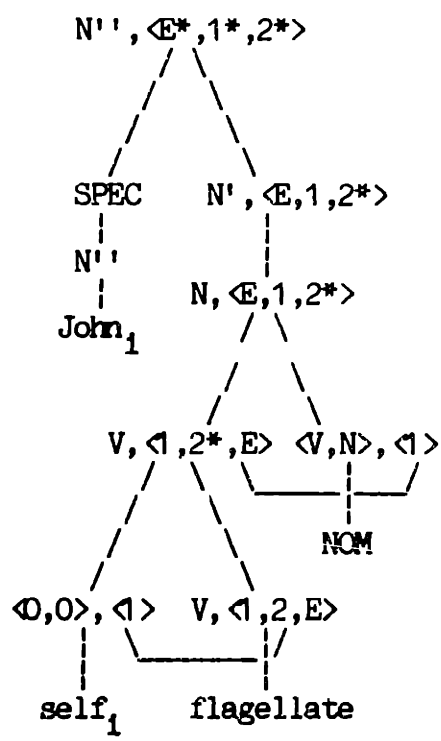
2.5.2 Self- in Derived Nominals.

We next turn to examples of the following form:

(226)
a. John's self-flagellation was a pitiful event to observe.
b. The aardvarks' self-annihilation was pleasing to the termites.
c. Self-destruction by Kamikaze pilots was often effective in achieving its goals.

I assume that the structure is as follows, and that self is bound by the subject (or agent) of the NP:

(227)



The semantic interpretation, simplifying using the equation given by the

coindexing is:

(228) v(e, N'') <<=>> Event(e) & [flagellate(John,John e)]

So, the NP refers to an e, where e is an event of John flagellating John.

What about cases like the following, where there is no overt subject or agent in the NP?

(229)
a. Self-flagellation is not a common practice these days.
b. Self-destruction is not a desirable end.
c. Self-castigation is often required for moral rectitude.

Binding appears to be local here since the meaning of (229a) for instance is that for one to flagellate oneself is not a common practice these days. In fact, it is obligatorily local as the following examples show:

(230)
a. They expected self-annihilation to be required.
b. They expected self-destruction not to be very nice.
c. The monks thought self-flagellation to be outdated.

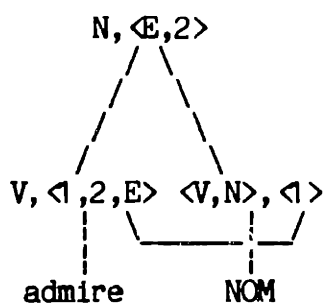
The crucial point about these examples is that although self can refer to the same set of individuals as specified by the matrix subject, it can only do so if the agent of the nominal is understood as being also coreferential with the matrix subject. This is different from the case of phrasal anaphors, which we discussed in Section 2.4.1.3, where it was observed that in many cases an anaphoric object of a nominal could be interpreted as coreferential with a matrix subject, even if the (understood) subject of the nominal was interpreted as being distinct in reference. This we

attributed to the fact that the External Theta Role is in general not necessarily discharged, although there seems to be variation, as also noted in the previous discussion, across speakers and across different example sentences.

But why cannot something of this form occur here? Why cannot the External Theta Role simply go undischarged in the self-nominal constructions, thus simultaneously allowing for free interpretation of the understood subject and for binding of the anaphor self outside the NP? So, why cannot (230b) mean that they anticipated that somebody or other destroying them would not be nice?

I think the answer lies in the correct interpretation of the notion of Complete Functional Complex. We assumed above (see the diagram in (175)), perhaps implicitly, that if an External Theta Role were to go undischarged in a derived nominal it would be undischarged across the category change from V to N. That is, the External Theta Role would be represented in the theta grid of the verb and not of the dominating noun:

(231)



Let us propose a term for this lack of discharge of the External Theta Role: we will say that the External Theta Role **hangs** across A and B, A and B being nodes in a tree T, if it is uninherited in a theta grid transfer (percolation) from A to B. In other words, to take a specific kind of example, if the External Theta Role is not inherited by a derived nominal along with the rest of the verb's theta-grid, then it will be said to hang. So, in the above structure, the External Theta Role hangs across V and N. If N has an anaphoric object, then, there will be no subject for the anaphor to make the NP count as a CFC for that anaphor and it will therefore be bound outside.

However, there is a difference with the prefix self in that self is not a sister of N but a sister of V dominated by N. In particular, V is the governor of self and also assigns it a theta role so the CFC for self will be defined in terms of V, not in terms of whatever N might inherit from V. Presumably, then, the External Theta Role of V is present in the grid (it has not **hung**) and is therefore taken to be part of the definition of the CFC for self insofar as self must be bound within a domain that satisfies the theta requirements expressed at the level of its theta assigner. If the External Theta Role must be satisfied as part of the definition of the CFC of self, then it cannot hang across V and N. So N will then inherit the External Theta Role and will therefore be obliged to discharge it in some way or other. This discharge will take place to the subject position of the NP, unless there is a by-phrase, in which case it will be discharged to

that. So, the presence of the subject will be forced and the binding will necessarily be local.

In the next section I consider one final set of cases which seem to support the analysis presented here.

2.5.3 Self- Prefixation to Nouns and Adjectives.

Self- also prefixes to clear cases of nouns and adjectives, such as the following:

(232)
a. A self-portrait was seen hanging in the Louvre.
b. Self-love can be objectionable.
c. John is really self-destructive.

There is presumably no doubt that self-portrait is an example of affixation of self to a noun. The other two cases are less obvious. So, why is self not attaching to the verb love in (232b) which is subsequently nominalized, and why is self- not attaching to destroy with subsequent affixation of -ive?

In fact, I think that love is not a case of affixation of NOM to a verb, but is, like other deverbal nouns which lack overt marking of their nominal status, simply a form which is listed as related to love; see section 2.4.2. for a discussion of such cases. This form is interesting in that it apparently has a theta grid like that of the verb love, as evidenced by examples of the form John's love of Mary, Monty's love of Lie Algebras, and

so forth. Nevertheless, compounds such as *woman love, *cat love and so forth are bad if interpreted as synthetic compounds: root compounds such as puppy love are okay. This makes sense if we assume, on the one hand, that whatever nominalization process is going on in verbs like love is marked for a particular verb, hence disallowing the nominalizer to be bracketed syntactically outside an $[N\ V]_V$ structure, and on the other that a nominal element cannot adjoin to the left of a deverbal noun such as love and be interpreted as a thematic argument since in order to be so interpreted, by the Uniformity Condition, Case would have to be assignable which it would not be; the left member of a noun-noun compound is presumably not a Case position. On the other hand, self-, being an affix, does not need Case and is therefore free to attach to such forms and be thematically interpreted. In fact, self- is generally affixable in forms such as these:

(233) self-love (*woman-love); self-hatred (*dog-hatred); self-knowledge (*algebra-knowledge)

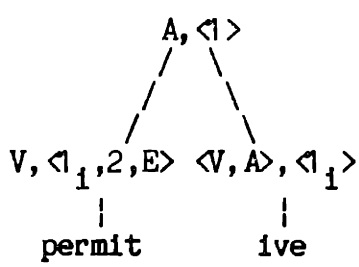
Turning now to the adjectival cases, there is very clear evidence that self- must be attaching outside the adjective in examples such as self-destructive. -Ive is a fairly productive suffix changing verbs into nouns. However, it is not a semantically transparent suffix; adjectives formed from verbs plus -ive are not predictably related to the verbal bases. So although destructive may well have as its referent the set of entities which have a propensity towards destroying, it is not the case

that constructive is related to construct in a similarly predictable way. Similarly, conducive is not obviously related to conduct. Of those that do seem to be related in an obvious way, only very few allow for complements:

(234)
a. (?) John is destructive of everything that he has tried to achieve.
b. *Mark is productive of work.
c. *Mary is not permissive of lewdness.
d. ?Eric is communicative of his ideas.

It seems that what we want to say about -ive adjectives is that (when they are related more or less transparently to the verb) they are formed by indexing a position in the theta grid of the affix with the external position in the grid of the verb. This will in general yield a structure of the following form, where the verb's thematic grid is not inherited by the adjective:

(235)

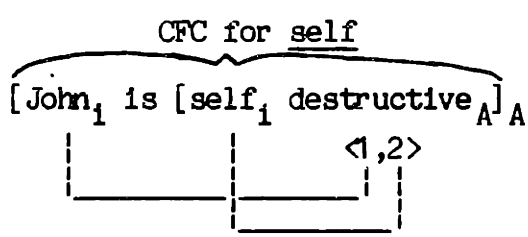


Destructive is simply marked in apparently allowing (though not requiring) that the thematic grid of the verb is inherited. However, we may fairly assume that in any event, the -ive deverbal adjective must be in general listed with particular verbs insofar as it does not normally have a

productive relationship with its base verb.

Given this, for forms like destructive, it is fairly clear that synthetic compounding will be impossible. Therefore, compounding will have to be with the adjective as a whole. This will rule out N-A compounds such as *city-destructive, again because the Uniformity Condition requires Case assignment, which is impossible in this case. Self- has no Case requirement, however, and is thus well-formed with such examples. It behaves perfectly normally in being discharged the internal theta role. The CFC will be defined as the entire complex in which there is a governor for self-, namely the adjective, and in which the theta grid of the adjective--in this case including the External Theta Role, which is the role assigned by the adjective to its modifiee--is satisfied. In a predicate adjectival construction this may be the entire sentence:

(236)



There is one other fact which is interesting, and which supports the idea that self- in its "compounds" is syntactically bound and not lexically bound. Assuming that nouns like portrait have an optional theta role which they can assign to the "object" of the portrait, then self can be argued to

be getting a theta role from the grid of portrait in the construction self-portrait. Yet, although it is clear that a self-portrait refers to a painting which one has done of oneself, there also seems to be a very strong desire on the part of speakers to have the self- bound by the subject of the NP. In fact, although it is perfectly obvious what such constructions would mean, speakers reject sentences of the following form:

(237) *The Museum of Modern Art's self-portrait by Picasso is hanging on display with the other cubist paintings.

There is a strong tendency to interpret the self as referring to the subject, which is absurd in this case, and which further means that the by-phrase is otiose. This seems to be a strong indication, then, that self- is syntactically bound, and not bound "in the lexicon."

This concludes the discussion of self- "compounds". I have argued that the distribution of these forms in the general case can be explained by the Binding Theory as construed in Chomsky (1984), with some auxiliary assumptions about the Case requirements for self-. Insofar as I have been successful, I have shown that there is no particular reason to claim, as some have (see Grimshaw (1983)) that there is a different kind of binding in lexical constructions from what we find in syntax. Certainly there appear to be differences on the surface insofar as local binding is always forced, which is not the case for phrasal anaphors. However, I have argued that this follows from the correct interpretation of the notion of Complete

Functional Complex.[33]

In the next section of this chapter I discuss extensions of the ideas introduced in the previous sections to slightly different kinds of morphology, including adjectival passives and deadjectival nominals.

2.6 Extensions of the Theory.

2.6.1 A Short Discussion of Adjectival Passives.

In this section I discuss some facts about adjectival passives in English which seem to confirm the ideas introduced in this chapter. I shall be adopting the excellent recent treatment of these forms by Levin and Rappaport (1985).

Levin and Rappaport assume that the formation of adjectival passives consists in two operations. The first part is the formation of a participle from the verb. The participle ending *-en* has the property that

---

33. Though Noam Chomsky has noted (p.c.) *John's self-portraits are not for sale* with the meaning that John has a collection of self-portraits and that his collection is not for sale. I happen to find this judgment rather dubious. Ken Hale has pointed out examples like *That is self-destructive.* where there seems to be no overt antecedent. I suggest that these are parallel to *Books about oneself are annoying*, where, again, there is no overt antecedent; most likely, there is an understood *one* which is the binder in such cases.

it suppresses the External Theta Role of the verb[34]. In syntax, following Jaeggli (1984), and Baker, Johnson and Roberts (in preparation), this suppression takes place under discharge via predication to _-en_, which is presumably in INFL. So in these cases it acts as an argument. In the lexicon, under my analysis this cannot happen since the internal theta role(s) will still remain undischarged at this stage. I assume then that the direct affixation of _-en_ to a verb simply allows the external theta role to hang in the sense introduced in the previous part of this chapter.

The second part of the formation of adjectival passives involves a rule which changes participles into adjectives. This is schematized by Levin and Rappaport as follows:

(238) _Rule of APF_-->$[V_{[Part]}]_A$

They suggest that the externalization of the internal argument is due to this conversion to an adjective. This is contrary to Williams' (1981) account where he explicitly states a rule of internal theta role externalization. They suggest, in fact, that the externalization of the

---

34. Though Fabb (1984) argues (pp. 49-51) that _-en_ behaves differently when governed by the auxiliary _have_ as in _I have killed the rutabaga_. He is concerned with Case assignment rather than with theta role assignment, and notes that while, under the LGB analysis, _-en_ absorbs Case in passives, it is cannot be said to do so in perfects. He suggests that the Case assigning potential of the verb is reinstated by the presense of _have_ which "balances out the [Case] features of the sentence." Something clearly also needs to be said about the theta properties of _-en_ to make this story work.

internal theta role(s) is simply the unmarked property of the conversion from verbs to adjectives[35].

One way to look at this process is to view thematic structure as being arranged somewhat like the layers of a small onion, with the external theta position being the outer layers, and internal positions occuping one or more inner layers. The suppression of the External Theta Role, as in the conversion from a verbal participle to an adjective could be viewed as peeling away the outer layer, and discarding it[36].

One important point which they demonstrate is that if a verb has more than one internal theta role, then any one may be the externalized role of the adjectival passive but all of of the obligatory ones must be satisfied (and all of the optional ones may be satisfied). This point is shown by

---

35. Evidence for this are forms with the affix _-able_ which clearly picks up the internal argument in some way: So 'This sentence is parsable' clearly means that this sentence is such that someone can parse it. Levin and Rappaport do not discuss this particular class of evidence for the assertion (though they do give other evidence), but it seems as though this property of internal role externalization is fairly common in very semantically transparent adjective formation.

36. In fact, given that the salient property of adjectival conversion is that of externalizing the internal theta role, it may be possible to explain _why_ _-en_ participles may become adjectives in the first place, as opposed to other participles, which generally do not. (On this, see Fabb (1984), p. 212.) Since _-en_ suppresses the External Theta Role, it has one of the salient properties of an adjectival affix. Given this, we might actually assume that the fact that participles in _-en_ can be adjectives follows from their ability to suppress the External Role, rather than the other way around. This is a somewhat speculative at the moment, however.

the following (based on 41, 42 in Levin and Rappaport):

(239)
a. cram the freezer: the freezer remained crammed
b. *cram the food: *the food remained crammed
c. cram the food into the freezer: the food remained crammed into the freezer
d. stuff the pillow: the pillow remained stuffed
e. *stuff the feathers: *the feathers remained stuffed
f. stuff the feathers into the pillow: the feathers remained stuffed into the pillow

In the remainder of this section, I discuss a couple of properties of adjectival passives which fall out of a combination of my analysis with that of Levin and Rappaport.

First of all, constructions of the following form are bad in English, with the reading that the left hand member is an internal argument of the verb:

(240) *chicken-killed, *dog-seen, *house-painted *baby-fed....

This is presumably because the internal theta role of the verb would be absorbed by the left-hand nominal, and there would be no theta role for the adjective to inherit. Of course, where there are two internal theta roles, such constructions would still be ruled out, this time by the Projection Principle, as discussed in section 2.3.5.3. So, *freezer-crammed food or *The food remained freezer-crammed are out for the same reason as *shelf-putter of books is out.

Interestingly, Lardil contrasts with English in this respect; assignment

of the internal theta role is possible in seemingly lexical passives, and shows up in an interesting way in Agent nominals in that language (thanks to Ken Hale for pointing out these examples and for providing the data, and see also Hale et. al. (1981)). The affix -n in Lardil is the equivalent of English (-er). It shows up in constructions such as:

(241)
dangka-nee -n
person kill -er
(= someone who kills people)

Constructions such as this one are presumably exactly like their English counterparts.

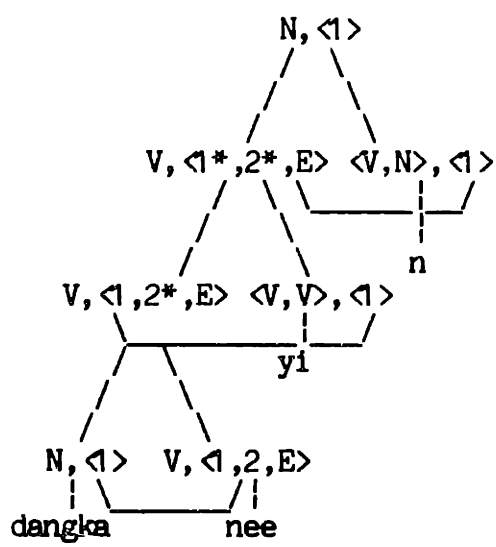
However, the passive morpheme -yi may also appear in these constructions, between the verb and -n, and when it does, it forms an instrumental noun:

(242)
dangka-ne -yi -n
person-kill passive -er
(= something with which to kill people)

There is a ready explanation for this behavior, given the observation that -yi in Lardil, unlike -en in English, never forms adjectives. What in fact seems to be happening in forms such as (242) is that the internal theta role is assigned to the noun which is sister to the V; the passive affix is then attached, but in this case we can actually discharge the External Theta Role to it, since the internal one is already discharged (ignoring,

again, the Event, which will be discharged in the change from V to N.) Since there is no V to A conversion in Lardil, there will be no requirement for an internal theta role to be around. Finally, when -n attaches, there is no External Theta Role for it to be attached to, and we may assume therefore that it defaults somehow to Instrumental, which seems to be the second most common interpretation of -er in English, following the Agent interpretation. One possibility, in fact, is that it is so interpreted under identification with the Event position. The construction will be as follows:

(243)



The interpretation of this, given that -yi is interpreted existentially, and the linking between -n and the event has some appropriate notation such as with(e,n), would be as follows:

(244) $v(\underline{z},N) \leftrightarrow E(\underline{e})E(\underline{x})[person(\underline{y}) \& kill(\underline{x},\underline{y},\underline{e})] \& with(\underline{e},\underline{z})$

Finally, note that expressed Agents in compounds with passive participles in English, such as man made are necessarily semantic arguments, as argued by Lieber (1983). This follows since attaching a nominal to a verb inside the -en cannot result in the nominal's receiving the External Theta Role, as argued in previous sections, and attaching it outside the participle will mean that it can only be interpreted semantically as the agent, since the External Argument is suppressed. This fact has an interesting consequence for self- "compounds" with participles. It will mean that the anaphor is not being assigned a theta role from a grid of an argument taker and that therefore it will not have a CFC. This should predict, given that the Binding Theory only applies to anaphors within CFCs, that the anaphor will not need to be bound syntactically. This seems to be right for such forms. So in self-addressed means that the Agent of the addressing addressed something or other to himself. But this Agent need not be syntactically present as in A self-addressed envelope is required. Similarly in Dutch (Eric Reuland, p.c.), zelf-gerepareerde auto (= self-repaired car) can refer to a car which is repaired by the owner himself, as opposed to being taken to a garage. But the Agent of the repairing does not need to be syntactically present. With this, I conclude this section adjectival passives. I turn next to a short discussion of deadjectival nominals.

2.6.2 On Deadjectival Nominals.

In this section I take a quick look at one class of deadjectival nominals in English, namely those formed with the suffix _-ity_. Even though I argued in the last chapter that a large class of _-ity_ nouns are listed, insofar as it is not predictably the case that just any latinate adjective can take the suffix, it does nevertheless seem to be the case that the constructions are semantically predictable. This is especially clear, perhaps unsurprisingly, in cases where _-ity_ is clearly syntactically productive anyway, namely the deverbal adjectives in _-able_. I shall assume henceforth that _-ity_ forms are semantically predictable in the general case, and that the necessity of listing a large class of adjectives is due to a restriction at some other level of the grammar, perhaps at the syntactic level, as in fact suggested in the analysis in the last chapter.

In fact it seems that there is good reason to argue that _-ity_, like NOM or _-er_ is an affix which allows inheritance of its sister's thematic grid. To see this, consider the following examples:

(245)
a. This sentence is parsable.
b. The parsability of this sentence is indisputable.

In (245a) the predicate adjective _parsable_ is assigning a theta role--ultimately the internal theta role of the verb _parse_--to the subject of the sentence. In (245b) it is precisely the same theta role which is

being assigned to the complement of the noun parsability. So, somehow or other, the internal theta role of parse is being inherited first by the adjective parsable and subsequently by the noun parsability.

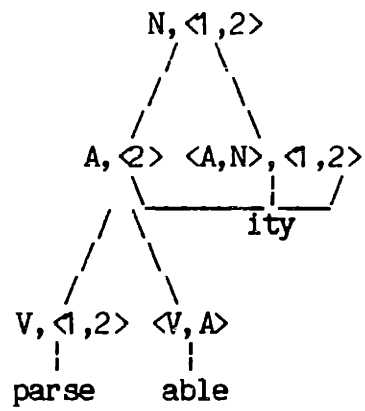
But what does the NP parsability of this sentence refer to? Clearly what it refers to is some state of affairs, in particular the state of affairs of some particular sentence's being parsable; what (245b) is claiming as indisputable is this state of affairs. Somehow this notion of state of affairs is introduced by the affix -ity. Let us suggest then that the full entry for -ity be given as follows:

(246)

ITY' = <ITY$_{\langle A,N\rangle}$ <1,2>, ity>
suf(ity)
State-of-affairs(1,2)

Thus -ity has two theta roles, one which is the External Role of the affix (and ultimately of the noun) and which refers to a state of affairs, and the other which will link under identification with the External Theta role of the adjective to which -ity attaches, and which expresses the particular kind of state of affiars denoted by the adjective. A word like parsability will have the following form, assuming (as discussed in a previous section) that the adjectival affix -able has the effect of hanging the External Theta Role of the verb and externalizing (one of the) internal arguments:

(247)



The interpretation of this form should be:

(248) v(<x,y>,N) <<=>> E(x)E(y)[**state-of-affairs**(x, parsable(y))]

So, parsability refers to some x, where x is the state of affairs of y's being parsable, for some y. A full NP such as the parsability of this sentence would be interpeted as follows:

(249) v(x,NP) <<=>> the x [**state-of-affairs**(x, parsable(this sentence))]

There is a certain parallel here with derived nominals insofar as derived nominals refer to an Event of some parties' performance of the action of the associated verb, whereas deadjectival forms in -ity refer to states of affairs of some entity or other's having the attribute of the associated adjective[37]. The difference is that the Event of derived

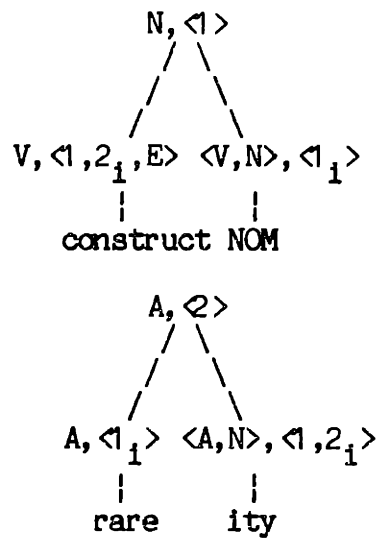
---

37. This fact was noted by Jespersen (1909-49), volume 4. p. 85.

nominals comes from the verb, whereas the state of affairs interpretation of -ity adjectives comes from the affix, and not from the adjective.

Also, there is another parallel between -ity forms and derived nominals, and that is that some -ity forms optionally have the equivalent of the result interpretation of certain specific nominals. The forms to which I am referring are the idiosyncratically interpreted words such as rarity with the meaning of "something which is rare" rather than "the state of being rare." Rarity in this interpretation refers to an entity which has the property of being rare. Similarly, construction in the result interpretation refers to an entity which has the property of being constructed. In the case of the derived nominal, I assume that NOM is optionally able to be indexed with the internal theta role of the verb, whose theta grid is therefore not percolated and remains locked up within the morphology. Similarly, -ity may be marked in certain cases to index one of its theta roles to the theta role of the adjective. If we assume that this is the 2 role, then the External Role (the 1) might just hang. This would yield structures, for the two forms, as follows:

(250)



2.7 Conclusion: Overview and Prospectus.

In this chapter I have given an account of syntactically productive and semantically predictable morphology. I have made a number of assumptions about the way various components of the grammar work, and I have also made assumptions about the form of the lexical entries of the various affixes; most of these assumptions I feel have been sufficiently justified.

The basic thrust of what I have tried to argue is that various components of the grammar, such as Binding Theory, Case Theory, Theta Theory, and so on, play a large role in determining well-formedness of supposedly lexical constructions; furthermore, and this is perhaps the more surprising point one needs to make few assumptions about the way these

sets of principles operate in "lexical" constructions in order to derive the right properties. So, I have suggested that Case Theory and Theta Theory, applying apparently as they do in the syntax, largely determine well-formedness of synthetic compounds, and the Binding Theory, again applying as it does in the syntax, determines the well-formedness of constructions containing the "lexical" anaphor self.

Of course, certain mechanisms have been argued to apply in collocations of lexical items (stems and affixes) which are subphrasal, but not in the corresponding phrasal constructions; it is certainly the case, for instance, as we have seen in the discussion of synthetic compounding, that theta discharge from a verb to its argument is not via the same mechanism--namely theta marking--as it is in VPs, but rather via the mechanism of theta identification. We derived this from the assumption that theta marking occurs only in the domain of maximal projections, a particular notion being given of what it means to be "in the domain of a maximal projection." So certain central guiding assumptions have proved necessary even though the components of the grammar themselves behave otherwise the same. As I pointed out in the last chapter, one could refer to these assumptions about the behavior of submaximal projections as being "the lexicon" or, to use a rather trendy term, "part of the alphabet of morphology." But to do so would have little theoretical significance, I feel; one could as easily claim that NPs constitute a separate component of the grammar because it is clear that nominals behave differently from verbs

in all sorts of salient respects. But there is little sense in this notion, I suspect, insofar as genuine components of the grammar, such as syntax and phonology, really do have properties which, I assume, are stated in terms of two markedly different vocabularies. None of this should be taken to mean that words do not exist. This no more follows from what I have said than that the (presumably correct) assertion that there is no "NP-component" would imply that there are no NPs. All I am claiming is that there does not seem to be a reason for positing a separate component of word formation.

Needless to say, these ideas are in direct confrontation of much recent work in morphology which argues that even productive morphology of the kind that I have studied in this chapter is governed by particular morphological or lexical principles of well-formedness. Many such well-formedness conditions have been proposed and I have no intention of giving an exhaustive discussion of these; the interested reader is referred to work by Williams (1981, in progress), Randall (1982, 1984), Grimshaw (1983), Grimshaw and Mester (1985), and many others. What I will do here is compare my approach with one recent proposal for a morphological principle, namely the Thematic Inheritance Principle of Randall (1982, 1984). Since we have already discussed this principle elsewhere in the chapter, this discussion is by way of review.

The Thematic Inheritance Principle is repeated again below:

(251) Thematic Inheritance Principle (Randall, 1984)
A category-changing operation which blocks the assignment of a θ-role blocks the assignment of all θ-roles lower on the θ-hierarchy.

where, θ-hierarchy is given as: Theme>Agent>Instrument, Source....

In previous sections of the chapter we have seen how two effects of the Thematic Inheritance Principle are in fact derivative. Firstly, Agent nominals in -er, which block assignment of the Agent insofar as they are the Agent, also block Instruments, and other obliques. This, we argued, was derivative of the assumption that such obliques are really modifiers of Events, an assumption which is based on facts about the semantics of sentences, and goes back to Davidson (1966). Since nouns do not have Event positions (although nouns may refer to events), such modification is predicted to be impossible in Agent nominals anyway. Derived nominals do have Event positions insofar as they refer to events, so we would predict that Event modifying obliques, such as Instrumentals, should be possible here. This is correct apparently, since all speakers I have consulted have no trouble with constructions of this form; this is contrary to Randall who takes such forms to be ungrammatical.

The explanation of the absence of Event modifiers in Agent nominals extends to cover their absence in adjectives, under the assumption that adjectives, like nouns, do not canonically have event positions. The following are thus predicted bad, in agreement with Randall:

(252)
a. The plane is flyable (*to France)
b. The plane unflown (*into the wind)

There is one other class of cases which Randall discusses, and that is where a noun or adjective inherits the Theme from the verb. It is predicted by the Thematic Inheritance Principle that the Agent should be unavailable. In the case of nouns thi. is clear:

(253)
a. **an employee by John
b. **a destroyee by the Romans

This fact is predictable under my account under either of two assumptions. Suppose first of all that -ee is not semantically productive in which case it is simply theta indexed, under my approach, with the Internal Argument. In that case, there will simply be no inheritance of the thematic grid and the Agent should therefore be inexpressible. Now suppose that there is a thematic grid inheritance, and that -ee picks up the Internal Theta Role, presumably under theta identification. That will mean that the internal role is discharged via binding from Spec since it is now expressed as the head of NP. But if the Agent Phrase were internal to NP, as in a by-phrase, it would mean that the External Theta Role is being discharged before the Internal one. But this is contrary to definition and is thus impossible.

The case with derived adjectives is less clear. The Thematic Inheritance Principle predicts them bad with Agents, as does the idea, discussed here (from Levin and Rappaport (1985)) that in adjectives the

External Theta Role is simply dispensed with. But this seems not always to be so clearly true, as the following well-formed examples suggest:

(254)
a. This grammar is unlearnable by children. (example from Roeper, 1984)
b. The house remained surrounded by trees.
c. It's certainly uncontaminated by any cheese.

On the other hand, the following are clearly bad, as Randall notes:

(255)
a. *The kite is flyable by experts.
b. ??The kite was unflown by experts.

So this suggests that an alternative approach, such as that of Williams (1981) for the passive, which has the effect of internalizing the External Argument, cannot be right either, since it predicts that in such cases the Agent ought to be generally available. I do not know what the answer here is, but as Levin and Rappaport (1985) point out, "[t]he conditions which govern when the agent argument can and cannot be expressed are subtle."[38]

So we have seen that my approach fares just as well with the facts as does Randall's, without positing a separate principle of morphology to do so. Rather, the facts fall out from general principles, coupled with particular lexical entries for affixes.

---

38. Ken Hale has suggested that only real Agents are blocked: (i) The house was unsurrounded by trees, versus (ii) *The house was unsurrounded by troops.

What kinds of semantically transparent, theta grid inheriting morphology are likely to be present in the world's languages? I would predict that, as in English, theta grids should be inheritable by Agent nominals which will be able to thematically mark Internal Arguments, but will not be able to take Event modifiers. Event nominals will essentially inherit the entire theta grid of the verb, though they will be constrained differently from the verb, by virtue of being nouns, with regard to the discharge of the positions in the grid. Finally, a nominal referring to an Internal Role would be able to inherit the entire Internal part of the grid, but not the External Role for reasons just discussed; the External Role will hang across the change from N to V. These predictions obviously need to be tested cross-linguistically. The little work that I have done on the syntax of nominalizations in other languages suggests that they are correct but clearly more work needs to be done.

But what of theta indexing, the process by which an affix can be indexed with a position in an unprojected theta grid? Can there be a theory of this process? I have made use of this notation to mark what I have termed semantically untransparent morphology, i.e., that morphology which does not allow theta grids to be inherited. I suspect that a theory of this process is possible, although such a theory would have to be able to explain somehow when one is allowed to pick some particular position in the grid to index, and when some other position must be picked. It will have to account for the fact that, in the following examples, the External Theta

Role (256a), the Event (256b) and the Internal Theta Role (256c) are picked up:

(256)
a. a bore, a cheat, a thief(?)
b. a kick, a run, a walk
c. a kill, a command, an order

What I have studied here is a class of particularly productive and semantically predictable affixation processes. Of course, there is a lot more to productive morphology than that. So what of verbal affixes such as re-, which has interesting properties (see Randall (1982), and others), and what of the particularly productive though not semantically transparent process by which verbs can be zero-derived from nouns (see Lieber (1980), Kiparsky (1983), and others)? And what of root compounding which is often interpreted based on pragmatic information, but which is nonetheless completely productive in a language like English (see Downing, (1977)). I suspect that much productive morphology will be shown to be reducible in its properties to principles of grammar needed independently. Nevertheless, the semantics of some perfectly productive processes such as root compounding will undoubtedly evade any attempt to reduce them to linguistic constraints, narrowly construed. Root compounding and zero-derivation of verbs from nouns seems to be governed as much by real-world knowledge as by anything obviously linguistic, and they are thus not only extra-syntactic, but they are extra-linguistic as well. But apart from these possibly intractable cases, I suspect that many of the

principles apparently necessary in the government of well-formedness of lexical forms will be reducible to more general principles of grammatical organization. This may well turn out to be a wild claim, but I feel that it makes for an interesting research project. What I have done in this chapter is attempt to show that this project shows some promise for success.

Chapter 3

Anaphoric Islandhood

3.1 Introduction.

In this chapter I examine the Anaphoric Islandhood of lexical items. This phenomenon is illustrated by the following examples:

(1)
a. Drivers of trucks$_i$ fill them$_i$ up with Arco Diesel.
b. *Truck$_i$-drivers fill them$_i$ up with Arco Diesel.

In (1a), where trucks is an NP object of drivers, it is perfectly possible to interpret the pronoun them as being coreferential with trucks. In (1b), on the other hand, truck is contained within the compound truck-drivers, and it is impossible to understand the them as being coreferential with truck, except in some peripheral way. It seems in general, in fact, that parts of words cannot be referred to by pronouns; in fact, pronouns cannot occur inside words either:

(2) *Bill is a Reagan$_i$ite, but I'd never be a him$_i$ite.

Note too, that Anaphoric Islandhood of words holds of other languages besides English. The following examples from Chichewa illustrate this point:

(3)

a. *m- dya makoko$_i$ si- a- wa$_i$- kanda ndi jamu
Class eat husks not subj obj(them) like with jam
*'A husk$_i$-eater doesn't like them$_i$ with jam.'

a. wo- dya makoko$_i$ si- a- wa$_i$- kanda ndi jamu
rel eat husks not subj obj(them) like with jam
'An eater of husks$_i$ doesn't like them$_i$ with jam.'

Postal (1969) was the first researcher to discuss these facts within Generative Grammar. The main intention behind his paper was to uphold the theory of Generative Semantics which was developing at that time, in opposition to the theory of Interpretive Semantics which was under development by Chomsky and others at MIT. Within Generative Semantics two apparently unparallel forms such as pork and wombatmeat were derived from [MEAT] from [PIG] and [MEAT] from [WOMBAT] respectively. Both of them show Anaphoric Island effects:

(4)

a. *The best wombatmeat comes from young ones.
b. *The best pork comes from young ones.

Postal used this parallelism to construct a rather contrived argument that Generative Semantics could account for the data in (4) by showing both

cases to arise from parallel underlying structures. In the remaining discussion, I will, needless to say, not be concerned with examples like (4b). There is no reason to suppose that pork is underlyingly morphologically complex, and its Anaphoric Island properties will therefore be a trivial consequence of this fact (see also Fodor et. al., (1980)), for arguments against the sorts of definitions which Generative Semantics assumed underlay many apparently monomorphemic words.) It is examples such as (4a) in which we shall be interested.

This phenomenon has also been taken of late to provide evidence for the independence of syntax and morphology. Since syntactic processes such as pronominal coreference apparently cannot "see" inside words, this seems to indicate that for some reason or other, words do not have any internal structure as far as the syntax is concerned. Yet words do have an internal structure, so it seems plausible to suggest that this structure was derived in a component of the grammar which was divorced from syntax in many salient respects.

I am proposing a theory of grammar in which there is no such divorce between word structure and syntax. Syntax, therefore, could in principle "see" inside words. The Anaphoric Islandhood effects must be derived from something other than a separation of components, therefore. In fact, I shall argue that it follows trivially from the fact that the morphology deals essentially exclusively in non-maximal projections, and given the

reasonable assumption that only referential expressions--i.e., maximal projections- can serve as antecedents to pronouns, Anaphoric Islandhood is to be expected.

In the next section of the chapter I discuss Simpson's (1983) analysis of Anaphoric Islandhood, which is a fairly clear example of a theory which makes use of the supposed separation of syntax and morphology. I shall argue that although her theory certainly derives the facts, there is a conceptual problem with it which makes it seem implausible as the correct solution. In the following section I then present my own analysis.

3.2 Simpson (1983).

In developing a theory of the morphology and syntax of Warlpiri in LFG, Simpson suggests that one of the guiding principles should be the Lexical Integrity Hypothesis (LIH), which she gives in a revised version as follows (p. 75):

(5) Revised Lexical Integrity Hypothesis (Simpson, 1983)

> Constituent-structure processes (which include annotation of functional information, and indexing of anaphoric information) are blind to the internal structure of words.

From this principle, given the assumption in LFG that words are inserted into constituent structure fully formed, Simpson derives three properties

of lexical items.

First of all, she argues that the LIH prevents gapping from applying to parts of words:

(6)
a. John paid the electricity bills, and Mary the gas bills.
b. *John liked the play, and Mary dis- it

In the (a) example the second occurence of the verb paid is gapped out, leaving a perfectly grammatical construction. On the other hand, in the (b) example, it is not possible to gap the second occurence of like from the morphological construction dislike.

Nevertheless, as Simpson points out in a footnote, this issue is by no means clear cut insofar as there are cases where gapping does seem to be possible, with Level II affixes:

(7)
a. John is father- and motherless.
b. ?John is father- and Mary motherless

though parallel constructions with other lexical items are bad:

(8)
a. *John is hope- and careless.
b. *John is hope- and Mary careless.

Simpson suggests that in the cases where the examples are good, the argument structure is relatively transparent. So, in (7), fatherless and motherless both have fairly semantically transparent meanings insofar as they both mean roughly lacking an X, where X is father or mother. Hopeless

and careless have somewhat less transparent meanings, on the other hand. Apparently "concreteness (father as opposed to hope) is important for interpreting something as an argument." Simpson suggests that "giving up the Lexical Integrity Hypothesis for the sake of this restricted set of examples seems...unwarranted." Nevertheless, it is obviously the case that the LIH is at the very least not as strong a condition as its formulation would lead one to believe; its application seems to be affected by issues such as semantic transparency[1]. In fact, it may well be the case that Simpson has hit the nail on the head as far as finding the correct condition is concerned: it might well be that a combination of semantic and prosodic considerations is what is really at work in governing gapping[2,3]. In fact, it is well-known that semantic factors are at work in governing gapping in syntactic applications. For instance, identity of meaning is generally assumed between a gapped element and its licenser. Hence, the

---

1. Booij (1983), argues that similar examples of coordination reduction in Dutch and German, are subject to conditions of prosodic structure, which adds another dimension to the possible space of conditioning factors for this process.

2. See the previous footnote. Note in this regard that the fact that gapping and conjunction reduction is bad when Stratum I affixes are involved may have to do with the fact that Stratum I affixes are stress determining and thus affect the prosodic structure of the lexical items to which they attach.

3. As Ken Hale has pointed out to me, Spanish coordination-reduction facts such as clara- y distintamente ('clear(ly) and distinctly') are in line with this observation. Not only are the X-mente formations in question phonologically transparent, but they are also semantically transparent.

following are bad:

(9) *John killed a fly and Mary a conversation.
(10) *Hortense shut her mouth and Edna the door.
(11) *Eric bagged two deer and Freda the groceries.
(12) *Bill pulled the rope and Fred a muscle.

In each of these examples two meanings of the same verb, one of them somewhat more idiomatic than the other, occur in the two conjoined halves of the sentence. It is not possible to gap the second occurence of the verb, since the result of doing this is apparently ungrammatical, and is only possible if a humorous effect is desired. This is entirely analogous to the difference between (7) and (8). In father- and motherless the suffix -less has the same meaning (i.e., 'without') in both halves of the conjunction. In *hope- and careless this is not true; both hopeless and careless are idiomatic in the sense that both mean something more specific than 'being without hope' and 'being without care.' Hopeless in particular rarely means 'without hope' but is rather used to refer to something or someone that is beyond help. So it seems as though the grammatical examples that Simpson mentions really might serve as counterevidence to the LIH, and rather point the way to a more adequate interpretation of these data.

A second class of cases which Simpson considers as derived from the LIH are cases of inflectional morphology, such as Tense, Case and Number markers. In LFG this kind of morphology, as every other kind of morphology, is done in the lexicon. In fact, the LIH would prevent things

from being otherwise. However, it does not seem that consequences favoring LFG arise from this decision[4].

Finally, Simpson argues that the Anaphoric Islandhood of words is derivable from the LIH. Assuming that coreference between pronouns and antecedents, for instance, is notated by some sort of indexing in constituent structure, and given that this kind of coindexing is considered as a relevant syntactic process by the LIH, then it will follow that indexings such as the following will be illicit:

(13)

a. *Truck$_i$-drivers fill them$_i$ up with diesel.
b. *Wombat$_i$-meat is best from the young ones$_i$.
c. *Reagan$_i$-haters would never be seen standing next to him$_i$.
d. *Bill is a McCarthy$_i$ite and Fred is also a him$_i$-ite.

Is the LIH an independent principle of grammar? The answer, according to Simpson, is in fact "no;" she argues, following Pesetsky (1979), who

---

4. Ken Hale (p.c.) has pointed out an example from Papago which might actually be problematic for such a view:

Ṣaacu$_i$ ʔapt wuḍ o [e]$_i$-k
what Aux-2psg equational-particle fut stative-pred.marker
'What will you be?'

In this example the -k is a stative predicate marker which is construed with ṣaacu 'what.' However, it is stranded by the operation of wh-movement, suggesting that the lexical unit formed by ṣaacu+k is not an island for syntactic processes

originally made the suggestion, and Mohanan (1982), that the LIH is derivable from the Bracketing Erasure Convention (BEC). Remember that in Lexical Phonology and Morphology the BEC is supposed to apply at the end of every stratum in the lexicon erasing internal brackets; its effect is to make the morphological structure of a word derived at Stratum n invisible to phonological and morphological operations which apply at Stratum n+1. Given that the BEC also applies at the end of the last Stratum (following, in particular, Mohanan), it will follow that the internal structure of a word is invisible to syntax. In particular, as far as the syntax is concerned, the structure of truck driver will be the unanalyzable [truckdriver]. Since there are no internal brackets, there is simply no way to index the relevant part of the word.

As appealing as this idea might seem, however, it suffers from a conceptual problem. The BEC, as was argued also in Chapter I, is motivated (if at all) on phonological grounds. It seems strange, then, that phonological considerations should have an effect on syntactic operations such as indexing; whoever would have thought that coindexation is done on phonological strings? Furthermore, if as Kiparsky (1983) suggests, and as Hargus (1985) also argues, the BEC may be exceptionable with certain affixes, the question is raised as to whether or not the reduction of the LIH to the BEC is even feasible.

Needless to say, in a theory of morphology such as the one being

presented in this dissertation, the reduction of the LIH to the BEC is not even possible. In particular, even if Bracketing Erasure is motivated for the phonological component of the grammar, this will have no effect whatsoever on the syntactic representation of words insofar as the syntax will still be able to "see" the fact that truck driver is morphologically complex. We must therefore seek another way in which to derive the Anaphoric Island effects. We turn to this question immediately.

3.3 On the Non-referentiality of Subparts of Words.

3.3.1 Pronouns and NP Anaphors.

Why are constructions of the following form bad?

(14)
a. *[a picture of it_i]$_i$

Such forms are out in because the pronoun must be disjoint in reference with the containing NP, and it has often been assumed that there is either a grammatical condition (such as the i-within-i condition, (Chomsky, 1981)), or some sort of condition on semantic interpretation, which rules them out. So, for some reason or other it cannot be non-disjoint with the NP which dominates it. But why can it not be non-disjoint with the head of the construction, namely picture?

(15) *[a picture$_i$ of it$_i$]

The reason for this has to do with a principle which has usually been assumed but never formally stated, to my knowledge, and that is that anaphoric or pronominal NPs must have arguments--i.e. NPs--for antecedents. Picture is a noun, not an NP, hence not an argument, and as such, it does not have a reference of its own. It cannot therefore be coindexed with it. Let us state this restriction as follows:

(16) Anaphoric/Pronominal Argument Indexing Condition.

> In a configuration ...**a**...**b**... (linear order irrelevant) where **a** is the antecedent of **b**, and **b** is an argument, then **a** must also be an argument.

Why are NPs--i.e. arguments--the only referential nominal expressions? Note (following Higginbotham) that a noun such as dog has an open position which will eventually be discharged via binding from SPEC, or under predication to some other NP. 'Dog , -V +N, <1> "denotes each of the various dogs" (Higginbotham, p. 23). It does not refer in the sense that it does not pick out any particular dog or any particular set of dogs. Only NPs have this property.

Given (16), the Anaphoric Islandhood of words follows trivially. Word formation deals exclusively in submaximal projections. So, we argued in the last chapter that the left member of synthetic compounds are maximally N'. Furthermore, it seems to be generally true that affixes select for

nodes of type X^0. Given this, the following examples will be ruled out:

(17)
a. *John is a Reagan$_i$-ite but I don't like him$_i$.
b. *Truck$_i$-drivers fill them$_i$ up with diesel.
c. *Aardvark$_i$-hunters rarely find them$_i$ on the Veld.

Note in particular that the noun Reagan apparently must be N^0 since it occurs in the word Reaganite. Despite the fact that the intended referent of Reagan here is obvious, it is still the case that (16) applies to the construction in (17a) to rule it out. (16) is therefore a syntactic condition which is not directly sensitive to semantic opacity.

Furthermore, given the reasonable assumption that pronouns such as he and anaphors such as himself obligatorily subtend NPs, the following will be ruled out on purely syntactic grounds:

(18)
a. *Fred is a Reagan$_i$ite but I am not a him$_i$-ite.
b. *Edna is an aarvark$_i$hunter but I am not a them$_i$-hunter.
c. *Truck$_i$-drivers are often hefty individuals and them$_i$-drivers are also quite often found in truck stops.

So an construction such as *them driver will be ruled out for the trivial reason that them is an NP, whereas the left-member of a compound cannot be a maximal projection.

The same argument can be applied to examples with the verbal anaphor do so; such cases were also discussed by Postal, and some of his examples follow:

(19)
a. *[Smoke]$_i$rs really shouldn't [do so]$_i$.
b. *People who hand[carve]$_i$ pipes compete with those who don't hand[do so]$_i$

Do so apparently refers to VP predicates, and--we might also fairly assume--subtends a VP also:

(20)
a. John looked at Mary and Bill did so too.
b. *John looked at Mary and Bill did so at Sally.
c. Eric hunts aardvarks and Hortense does so too.
d. *Eric hunts aardvarks and Hortense does so pigs.

Since VP predicates (or predicates in general in fact) do not occur within lexical forms such as compounds, it is therefore not surprising that do so is unable to refer back to a verb within a compound, or that do so cannot itself occur within a compound. This will, of course, require an extension of (16) to cover not just arguments but predicates as well:

(21) Anaphoric/Pronominal Argument/Predicate Indexing Condition.

In a configuration ...a...b... (linear order irrelevant) where **a** is the antecedent of **b**, and **b** is an argument/predicate, then **a** must also be an argument/predicate.

Note that this condition will extend to the following type of example,

pointed out to me by J. Higginbotham:

(22) Bill considers Joe [a fool]$_i$ but I don't consider him so_i

Clearly, since so is anaphoric for a fool, which is an NP predicate, the condition (21) applies to determine the well-formedness of the example.

We have thus been able to derive Anaphoric Island effects from a generalization about the indexing constraints on NP anaphors, a generalization which seems motivated for reasons independent of the considerations particular to the determination of well-formedness of "lexical" constructions. We now turn to some residual cases.

3.3.2 Non-Anaphoric-Islandhood with Submaximal Projections: Some Examples.

Note, that the Indexing Condition only applies to anaphors which are maximal projections. In the last chapter we argued that self- in compounds such as self-admirer was bound locally by the affix -er. Now, -er is presumably not a referential expression since it is not a maximal projection, but this is irrelevant since the Indexing Condition has nothing to say about this particular indexing, as self- is not an NP. Of course, self- may be coindexed with an NP antecedent, as in John's

self-destruction, but there is no requirement that it do so[5].

Also unconstrained by the Indexing Condition are examples of X^0 movement into morphological constructions; such phenomena are discussed extensively in Baker (1985c). One such example is Noun Incorporation, briefly mentioned in the last chapter, and investigated in Baker (1984; 1985c). Noun Incorporation creates structures of the following form:

(23) $\ldots.[_{VP}\ [_{V}\ldots N_i\text{-}V\ldots]\ \ [_{NP}\ [_{N}\ t_i]\ldots.]]\ldots.$

Again this type of structure is licit according to the Indexing Condition, since the trace, an anaphor to be sure, is nevertheless not a maximal projection and is therefore not constrained by the Indexing Condition.

As Baker notes (e Sadock, 1980) there are apparent violations of the Anaphoric Island condition in Noun Incorporation languages such as Greenlandic Eskimo:

---

5. A form like a Reagan$_i$ite's self$_i$-destruction will, of course, be ruled out simply by the fact that Reagan would not c-command self-:

$[_{VP}[_{NP}$ a $[_{N'}[[$Reagan$]$ ite$]]]$ $[_{N'}[$self-destruction$]]]$

(24)

Suulut timmisartu-lioq-poq. (pro) suluusa-qar- poq
Søren-abs airplane -make-3sgIndic it wings have 3sgIndic

aquute-qar- llunilu
rudder have 4sInf-and

"Søren made an airplane$_i$ (lit. airplane$_i$-made).
It$_i$ had wings and a rudder."

This fact is unproblematic, however, since we may fairly assume that what the pronoun is coindexed with is the NP subcategorized for by the verb, rather than directly with the incorporated N^0. So, the structure of the VP in the first sentence, according to Baker, is:

(25)

$[_{VP}\ [_{V}\text{timmisartu}_i\text{-lioq-poq}]\ [_{NP}\ [_{N}\ t_i]]]$
airplane make 3sgIndic

The pro in the following sentence is not coreferential with the incorporated noun timmisartu, but rather with the stranded NP.

One class of cases which might seem problematic for the Indexing Condition are the familiar cases of pro-drop in many languages. Here, at some level of abstraction we find structures such as the following:

(26) $[_{I''}\ [_{NP}\ \text{pro}]_i\ [_{I'}\ [_{I}\ldots\ \text{AGR}_i\ \ldots]\ldots$

AGR is a governor of the NP subject pro, since it is part of INFL, and it is also coindexed with pro; apparently an NP pronoun can be coindexed with a submaximal projection (assuming that is what AGR is) after all. The particular problem arises, of course, from the fact that in many languages, AGR merges with the verb at some point in the syntax to form a single lexical item. Thus, apparently coindexation of a pronoun with an "antecedent" within a lexical item is allowable.

But this kind of example is really irrelevant for the Indexing Condition since the latter explicitly constrains cases of antecedent coindexation of pronouns and anaphors, and the AGR/pro coindexation is surely not a coindexation of this kind. AGR is not an antecedent of pro. In fact, if it were, such constructions would surely be violations of Condition B of the Binding Theory, since a pronoun would be locally bound by an antecedent. Presumably, rather AGR licenses the existence of pro under coindexation. This, for example, is the position taken in McCloskey and Hale(1984), who suggest a condition such as the following (= their 82):

(27)

*pro unless governed by AGR
[αF] [αF]
where [αF] is some combination of person-number features.

This is a condition of licensing rather than antecedence, which makes the Indexing Condition irrelevant for such forms.

Needless to say, a similar way out of this apparent problem is necessary

in a lexical framework such as LFG. So, an affix which marks for third singular on verbs would be listed with the following equations:

(28)
(↑ SUBJ PRED) = 'PRO'
(↑ SUBJ NUM) = SG
(↑ SUBJ PERS) = 3

Of course, therefore, syntactic processes (which are sensitive to the subject of a pro-drop sentence in anybody's theory) can see inside a lexical item, if the notion of see is broadly construed.

Possibly problematic for the present analysis are cases involving the Identity of Sense pronominal one. One does form ungrammatical examples with compounds:

(29)
a. *John is a dog-lover and I am a one-lover too.
b. *Mary is a cat-fancier but I don't fancy one.

The reason that such examples might be problematic is that one apparently may be an N^0. So it may head NPs, as the following examples from Travis (1984) indicate:

(30)
a. I saw a picture of Debbie in the living room and one of Konrad in the dining room.

b. the picture of Julia and the one of Suzanne

Of course, one may also subtend an NP as in I saw a picture of Harry and you saw one too, so that might be the reason that examples like (29) are

bad. However, one can force the interpretation that one is submaximal by modifying one with an adjective. Such examples do not improve in grammaticality:

(31)
a. *I am a [black dog]-fancier but you are a [brown one]-fancier
b. *John loves the red Porsche but I am not a [red one]-lover

However, there is reason to assume that such constructions with one must always be construed as NPs. Thus, even though one in green ones is not an NP, the whole construction must be. The reason is simply that in order to be interpreted as a pronominal expression, forms with one must be arguments. Such a state of affairs would, of course, immediately rule out examples like (31) since here the one nominals could not be maximal projections, hence not arguments.

Independent evidence for this assertion comes from the fact that one nominals cannot be attributive. Consider examples such as the following:

(32) I have a white egg omelette but John has a brown egg omelette.

Here the N' white egg and brown egg are modifiers of the head noun omelette; the accentual pattern of both NPs, with the primary accent being on omelette indicates that these cannot be compounds. Similarly, white egg and brown egg are N', not compounds, again because of the accentual pattern (with primary stress on egg). However, while (33a) is possible, (33b) is definitely not:

(33)

a. I like a white egg but John prefers a brown one.
b. *I like a white egg omelette but John prefers a brown one omelette.

The constraint would appear to be that one nominals cannot occur in non-argument positions; i.e., in order to be able to be construed pronominally, they must be arguments. Modifiers are not arguments and are therefore not maximal projections (see Higginbotham, 1985b.) But this in turn means that they must be NPs, which precludes them from occuring within compounds. One nominals are therefore not a problem for the analysis of Anaphoric Islandhood proposed here.

3.3.3 When Maximal Projections Do Occur within "Lexical" Constructions

There are cases, nonetheless, where, for various reasons having to do with the syntactic properties of the particular language, full maximal projections may occur within "lexical" constructions. We saw some such cases from Dutch in the last Chapter:

(34)

a. de van de boom vallende appels et -er
the from the tree falling apples eat-er
"the eater of apples falling from the tree"

b. de in de oven gebakken cake et -er
the in the oven baked cake eat-er
"the eater of cakes baked in the oven"

c. de in de tuin groeinde tomaten bewonder-aar
the in the garden growing tomatoes admire -er
"the admirer of tomatoes growing in the garden"

Is it possible to refer to such NPs? Perhaps not surprisingly the answer is "no" (thanks to P. Muysken for this example):

(35)
*De van de boom$_i$ vallende appels et -er
the from the tree falling apples eat-er
"The eater of apples falling from the tree$_i$

keek naar de bladeren ervan$_i$
looked at the leaves from-it
looked at it$_i$'s leaves."

This might seem to be problematic, but I suspect that there is a straightforward explanation for this fact. One might suggest, for instance, that since compounds such as Agent nominal synthetic compounds refer to generic activities, it is impossible to force a definite interpretation of an NP within a compound, which is presumably what a pronominal like ervan would seem to require. On the other hand, however, the explanation might be purely structural; although the pronoun does not c-command its intended antecedent, it is nevertheless the case that the antecedent is quite deeply embedded compared to the pronoun. Notice, in fact, that the English translation of this example sounds distinctly odd, as do the following cases:

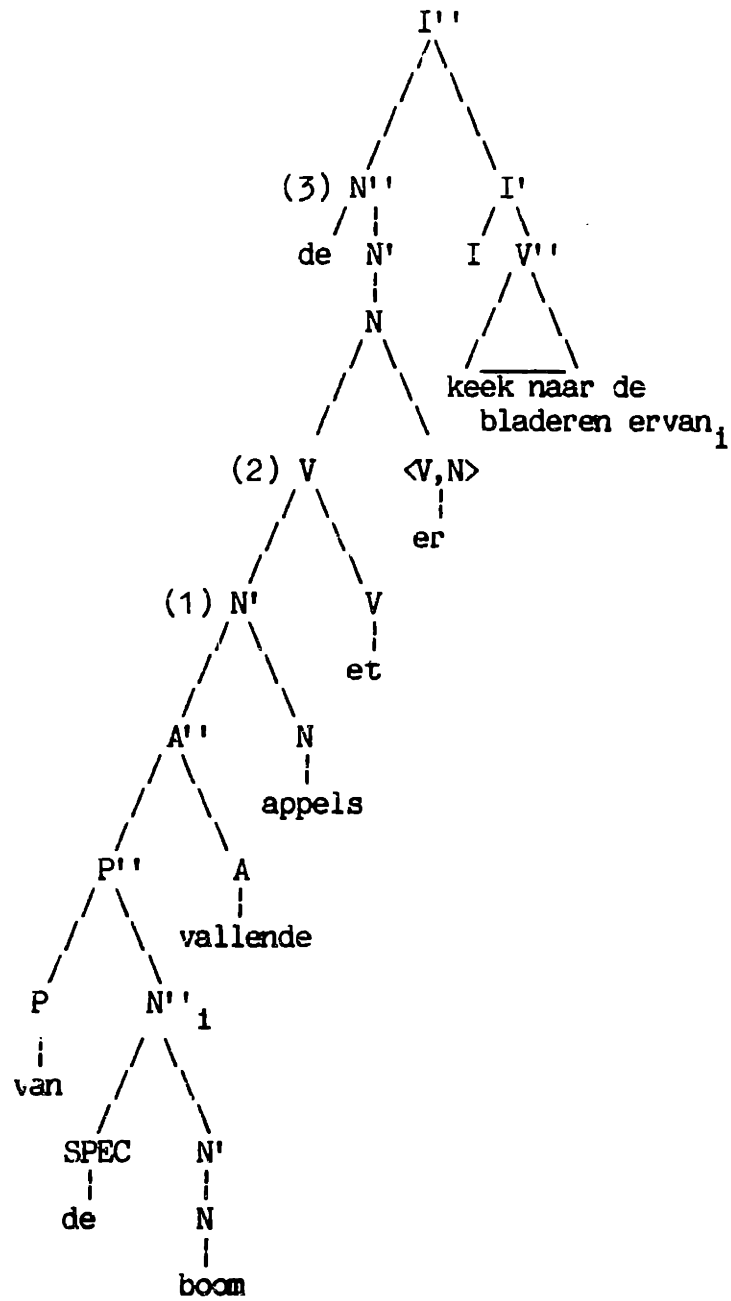
(36)

a. *Someone who likes dogs with hats$_i$ doesn't wear them$_i$.

b. ??People who drive trucks with large tires$_i$ don't like them$_i$.

In each of these cases, the intended antecedent has the same level of embedding as de boom does in the Dutch example. So, in (36a), hats is embedded (1) as a complement of the NP headed by dogs and is further embedded (2) under the relative clause I'' and finally (3) under the NP headed by someone. Similarly, de boom is embedded (1) under the N' headed by appels which is embedded (2) as the object of the verb eten which is embedded (3) further under the NP headed by eter:

(37)



In any event, whatever the precise constraint may be, it seems to be the case that it applies to "lexical" and non-lexical constructions alike.

Of course, if we could find exceptional lexical constructions where an NP does occur which is not too deeply embedded, then we would certainly expect to be able to get apparent violations Anaphoric Islandhood. Such cases occur apparently in Warlpiri (thanks to Mary Laughren and Ken Hale for the data in this discussion.) Warlpiri has a nominalizing suffix -ngu which attaches to verbs to form nouns. Apparently, with the exception of lexicalized forms such as karla-ngu (dig+-ngu) 'digging stick', such constructions are obligatorily in the form of synthetic compounds: so, wawirri-pu-ngu (kangaroo+kill+-ngu) 'kangaroo killer' is fine, but *pu-ngu 'killer' is not. NPs are allowed in these constructions as the complements of the verbal head which implies that pronominals ought to be fine within these forms, and be able to have antecedents which are also the left members of a -ngu compound[6]. All of this is evidenced by the following example:

---

6. We may assume that since the sister of V is in fact an NP, that the discharge of the verb's internal theta role is via theta marking rather than via theta identification. This will imply, in turn, that the projection of V within the compound is maximal, given the discussion in section 2.3.5.2 of the last chapter. Perhaps, then -ngu is <VP,N> subcategorizing for VPs and forming nouns. This might derive the fact that complements of the V are obligatory in the general case, since the Projection Principle would require their presence. However, the derivation would not be trivial insofar as Warlpiri allows empty pronouns as a normal syntactic expression of (definitely interpreted) objects; one would have to explain why such pronouns do not occur in synthetic compounds headed by -ngu.

(38)

Jampijinpa [[wawirri yalumpu-wardingki]$_i$ pu] -ngu.
kangaroo that place kill -er

?Japaljarri [nyanungu$_i$ pu] -ngu-yijala.
that kill-er also

"Jampijinpa is a killer of [kangaroos in that place]$_i$.
Japaljarri kills those$_i$ also."

The Warlpiri example is perhaps a little odd insofar as it is hard to interpret the left member of these compounds as being specific in reference, which is what a pronoun such as nyanungu implies. Other than that, these constructions are apparently well-formed, which is what is predicted by the analysis of Anaphoric Islandhood presented here.

3.4 A Summing Up.

In this short chapter I have presented an analysis of the long-observed Anaphoric Island phenomenon associated with "lexical" constructions. I have argued that the relevant data are explained by a condition on indexing of argument and predicate pronominals and anaphors--a condition which is arguably required for reasons independent of the question of the Anaphoric Islandhood of words. It seems, therefore, as if this property of words is not indicative of their being derived in a separate component of the grammar from the syntax, and hence internally "invisible" to syntactic

processes. Rather, syntactic processes may well "see" the internal structure of words, but fail to make use of this fact for the trivial reason that such constructions lack the right properties for coindexation.

Chapter 4

Concerning Lexical Phonology.

4.1 Introduction

In this chapter I turn to a discussion of the theory of Lexical Phonology and Morphology (LPM), which has been developed over the past few years by a number of researchers including Kiparsky (1983a,b), Mohanan (1982), Mohanan and Mohanan (1984), Halle and Mohanan (1985), and Pulleyblank (1983).

In this section I give a brief overview of the theory of LPM. I also show that, taken as a theory of the organization of the phonology, LPM is not necessarily incompatible with the approach taken in this thesis. In the second section, I argue that LPM is not a theory of word-formation; in particular, I shall attempt to show, contrary to the claims of Kiparsky and others, that stratum ordering does not constrain word formation per se, and that apparent stratum ordering effects can be reduced to other (independently needed) principles; this point is important for my approach

to morphology as I argue below in the section where these questions are addressed. Finally, in the third section, I shall examine LPM as a theory of the organization of the phonology. I shall in particular be examining issues such as Bracketing Erasure Cyclicity, and Structure Preservation. I shall suggest that LPM is not as useful a theory of phonological organization as it might at first seem and I shall sketch an alternative theory which, though less highly structured is just as explanatory as LPM. It will turn out that phonology applying "in the lexicon" is not obviously distinct in its behavior from post-lexical phonology, at least not as distinct as envisioned by LPM.

4.1.1 A Brief overview of LPM

In this chapter I shall be assuming some familiarity with LPM and I shall therefore not give an exhaustive review of the theory; many of the more substantive claims of the various versions of LPM will come up in subsequent discussion anyway. The following is only intended as a brief overview of the main points.

One of the problems with discussing LPM is that there is really no one theory of this model of phonology and morphology; there are at least two, namely that of Kiparsky (1983a,b 1984) and Mohanan (1982). What is generally agreed upon currently, however, is that the morphology is organized into a number of strata of word-formation. At each of these

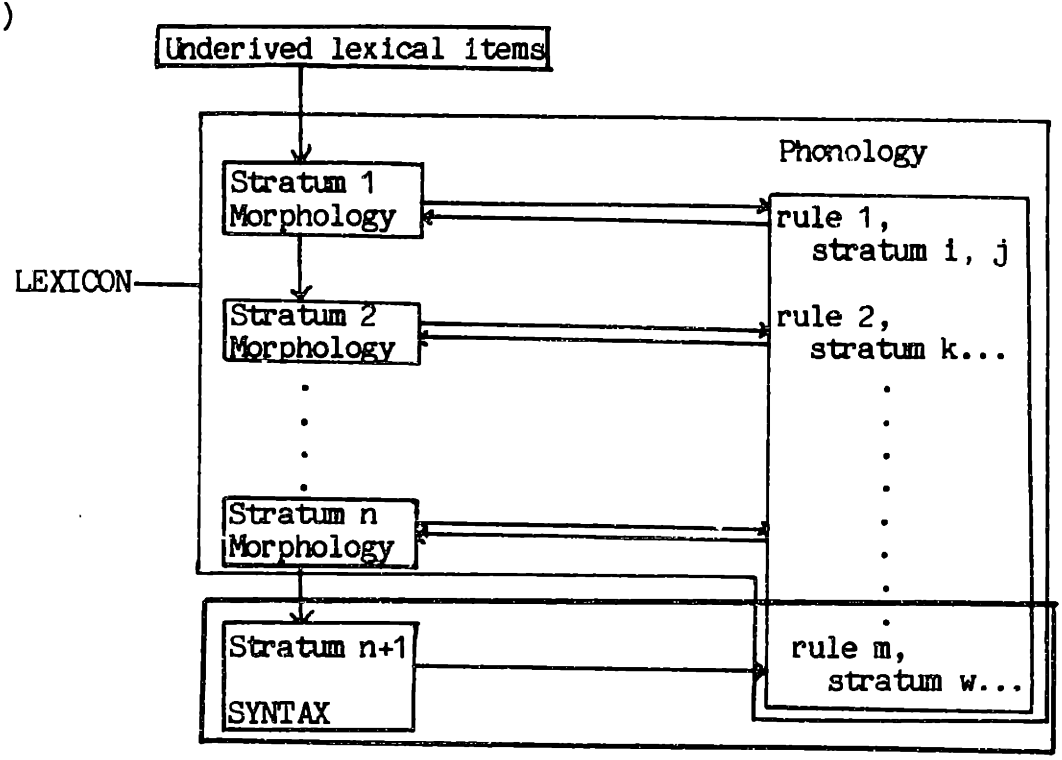
strata, certain affixation processes occur. At some point, either after each affixation at the given stratum or after all affixations which are to be done at that stratum, the derived form is submitted to the phonology where it is checked for rule application[1]. A phonological rule will apply to a morphologically derived form just in case two conditions are met; first of all, of course, the structural description of the rule must be satisfied, and secondly, it must be marked to apply at the particular stratum at which the morphological operation has applied. A diagram of this model is given below[2]:

---

1. The former would be a cyclic stratum, the latter a non-cyclic stratum. The idea that both kinds of strata may exist is proposed in Mohanan and Mohanan (1984) and Halle and Mohanan (1983). Kiparsky assumes that all strata are cyclic.

2. From Archangeli (1984a). This particular diagram actually illustrates a model with only cyclic strata (note the back and forth arrows between the morphological levels and the phonology.)

(1)



So, each rule is listed in the phonology as to the particular stratum at which it applies. Stratum n+1 in this diagram represents the post-lexical stratum; after a word has been formed in the lexicon, it is inserted into syntactic structure where it will be subjected to principles and rules of syntax, but also to further (post-lexical) rules of the phonology. What is the distinction between lexical and post-lexical rule applications? A fairly neat summary is given by Pulleyblank (1983):

(2)

LEXICAL	POST-LEXICAL
a. may refer to word-internal structure	a. cannot refer to word-internal structure
b. may not apply across words	b. may apply across words

c. may be cyclic	c. cannot be cyclic
d. if cyclic, then subject to strict cycle	d. non-cyclic, hence across-the board
e. structure-preserving	e. need not be structure-preserving
f. may have lexical exceptions	f. cannot have lexical exceptions
g. must precede all post-lexical rule applications	g. must follow all lexical rule applications

One point which needs to be clarified is what it means to be structure preserving. Lexical rules are claimed to be structure preserving in the sense that they do not add to the lexical segmental phonological inventory of a particular language (see, e.g., Pulleyblank, 1983, pp. 24-5); to be structure preserving, for instance, rules cannot introduce distinctive segments into the lexical phonology which are not lexically distinctive To give a simple example of this, no lexical rule of English phonology introduces aspiration on voiceless stops; such aspiration is not lexically distinctive. Post-lexical rules, however, need not be structure preserving according to this theory. Hence it is a post-lexical rule which is responsible for aspirating voiceless stops. That this is indeed a post-lexical rule is confirmed further by its other properties: it does not refer to word-internal structure, there is no reason to believe that it is cyclic, it has no lexical exceptions and it is known to follow at least one group of lexical rules, namely those rules assigning stress. We will have more to say about Structure Preservation later in the chapter.

There are a couple of other characteristics of the theory which we need to discuss here. One, which we have already seen mentioned in previous chapters is the Bracket Erasure Convention (BEC; also called the Opacity Principle (Mohanan, 1982).) The particular version quoted here is from Halle and Mohanan (1985):

(3) Bracket Erasure Convention

> After the application of all rules on a stratum, the brackets between the morphemes are deleted, so that reference to the constituent morphemes becomes impossible at subsequent strata.

I will discuss in a subsequent section then supposed support for this convention.

Finally, according to most researchers now, following originally Mohanan (1982), there is the stipulation that if a phonological rule R applies at both stratum m and stratum n. where n>m, then R must also apply on all strata m+1 through n-1. This is stated by Mohanan (p. 21) as the Stratum Domain Hypothesis (SDH):

(4) The domain of a rule is specified as a set of continuous strata.

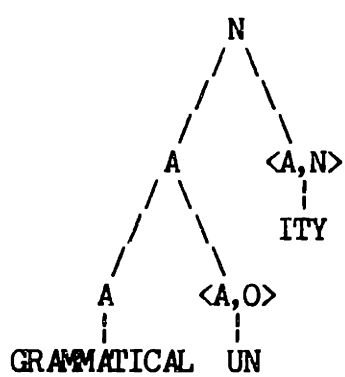
Having now given a brief overview of the generally-agreed-upon claims of lexical phonology, I now turn to a discussion of the compatibility of LPM, viewed as a theory of the organization of the phonology, with the approach to morphology being presented in this dissertation.

4.1.2 LPM as a Theory of Phonological Well-formedness

In previous chapters of this thesis, chapter two in particular, I have argued that syntactic well-formedness of a word can be dealt with in the syntax, by examining the properties of the particular morphemes which make up a word. For example, a word like parsability is well-formed because parse is a verb, able is an affix which takes verbs and makes adjectives, and ity is an affix which takes adjectives and makes nouns.

In Chapter 1, I argued that in the mapping from the S-structure representation of a word to the PF component, the Mapping Principle exerts a minimal requirement on the character of the phonological representation of the word. Take again a word like ungrammaticality, which I have taken to have the following syntactic representation (linear order irrelevant):

(5)



All the Mapping Principle requires (and the reader is referred to the relevant sections of Chapter 1 for extensive discussion,) is that un- be to

the left of grammatical and adjacent to it, and that -ity be adjacent to and on the right of the phonological representation of [GRAMMATICAL UN]. In the notation developed in Chapter 1, this set of constraints would be notated as follows: un^grammatical^ity.

The question that we are now interested in is how the phonology interprets such a string. That is, what structure (if any) would the stratum ordered phonology assign to this particular linear ordering of the phonological representations of these affixes?

Suppose, first of all, that the prefix and the suffix are to be added at the same stratum, according to the particular notation on the affixes; such would be the case, for example, with the word inability, which has two stratum I affixes. Then, assuming that only one phonological affixation takes place at a time, presumably either of the two following phonological bracketings is possible:

(6) [[in abil] ity], [in [abil ity]]

However, we might assume that there is a principle such as the following:

(7) Default Phonological Interpretation Principle (DPIP)

> When the phonology imposes no requirements of its own on the phonological bracketing of a string, pick an interpretation which yields a bracketing congruent to that of the

syntactic representation.[3]

This principle seems necessary, for instance, to account for the fact that in compounding, the phonological bracketing apparently mimics the syntactic bracketing. So city-dog house with the meaning of "a house for dogs which dwell in cities," has a different stress pattern from city doghouse (= "a house for dogs, which is in the city."):

(8)

1 4 3	1 3 4
city-dog house	city doghouse

This is derivable, according to most theories of stress assignment (e.g., Chomsky and Halle, (1968)), if we assume that in the former case, city and dog are phonologically joined together, followed by city-dog and house. The other compound will take the alternate route.

What, then, if the phonology does require that the phonological bracketing be non-congruent to the syntactic bracketing? For instance, take a classic Bracketing Paradox such as ungrammaticality. Here the suffix -ity must be phonologically inside the prefix un- for the simple reason that -ity is ordered at a earlier stratum (Stratum I) to un- (which

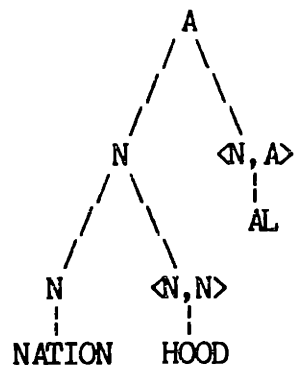
---

3. Marantz (1984b), suggests that the constraint is that phonological representations are canonically right-branching. This is an interesting suggestion insofar as it makes the correct prediction that in many languages the prefixes are phonologically external to the suffixes. I shall not be investigating Marantz's suggestion here.

is ordered at Stratum II). The string un^grammatical^ity which is provided by the Mapping Principle will thus be interpreted as follows: At Stratum I, -ity will be phonologically affixed to the stem grammatical and relevant rules of the phonology will be applied. Then Stratum I will be exited and Stratum II entered. Here, un- will be affixed to grammaticality and relevant phonological rules applied. All that is required in general, in fact, is that there be a phonological interpretation of the syntactic representation which is both consistent with the linear ordering imposed by the Mapping Principle and also consistent with the Stratum Ordering restrictions on the affixes (or if there are no such phonological restrictions, with the DPIP.) Clearly, a representation for ungrammaticality such as [un [grammatical ity]] is consistent with the representation provided by the Mapping Principle since it preserves the ordering in that representation. This point was, of course, extensively discussed in Chapter 1.

In addition to picking a well-formed phonological representation for a particular morpho-syntactic bracketing, Lexical Phonology, interpreted as a theory of phonological well-formedness, can rule out a particular construction if there is no such well-formed phonological representation. To take another previously discussed example, a non-word such as *nationhoodal would have the perfectly well-formed syntactic representation:

(9)



The Mapping Principle will yield the following minimal requirement on the phonological representation: nation^hood^al. And it will turn out, clearly, that there is no possible phonological interpretation of this string which will satisfy the requirements imposed both by the string and the stratum-ordered phonology. This is simply because in this case, unlike the case of ungrammaticality, there is no ambiguity as to which affix must be attached first. Clearly the order of phonological affixation must be -hood followed by -al. But this is ruled out by the fact that in order for the phonology to proceed in this fashion, we must enter Stratum I, where -al is attached, after we have applied rules at Stratum II, where -hood is attached. But since Stratum I is ordered before Stratum II, this is an impossibility. The structure will thus be ruled out in the PF component even though it is apparently well-formed from the point of view of the syntax.

Thus, if we allow that LPM is a model of the well-formedness of the phonological representation corresponding to syntactic constituent

structures, there is no inconsistency between that model and the one which I am proposing for morpho-syntactic well-formedness in this thesis.

On the other hand what is not possible is to interpret LPM as a theory of word-formation. In particular it cannot constrain the syntactic well-formedness of a word. Consider the following: Suppose that a syntactic representation [[A B] C], linear order irrelevant, is well-formed, and that the alternate bracketing [[A C] B] is ill-formed, where A is one of [+-N, +-V], and B and C are affixes. Suppose further that **a**, **b**, and **c** are the phonological representations of A, B and C respectively, and that **b** is a prefix and **c** is a suffix, and finally that **b** is stratum-orderd after **c**. This would yield the following phonological representation: [**b** [**a c**]]. This is, of course, a classic Bracketing Paradox. LPM would predict, as indeed it does for Bracketing Paradoxes in general, given that there are no extensions to the theory (such as the suggestion of Kiparsky discussed in Chapter 1), that such cases would always be ungrammatical. The only possible syntactic representation for [**b** [**a c**]] would be the congruent [B [A C]], which we stated above to be ill-formed. Thus, the stratum ordering in LPM can restrict the well-formedness of words which would be well-formed on purely syntactic grounds. In contrast, the theory I am proposing here would clearly claim that such cases will always be grammatical, unless of course, there are other principles (including, but not limited to, morphological blocking) which would rule out either or both representations. My basic claim is

that provided a well-formed phonological interpretation exists for a string which is provided by the Mapping Principle for a particular syntactic structure, then there is no general principle which would rule out the word; a word would only be ruled out if there were no such phonological structure. So LPM cannot be a theory of word-formation in the sense that it cannot, according to my theory, be allowed to strongly constrain the relationship between the syntactic and phonological representation of a word. A corollary of this is that there is no sense in which a morpho-syntactic operation, such as the general process in some language of forming words of one category from words of another category can be said to reside at any particular stratum of word-formation. Compounding in particular cannot be stratum ordered on my account: affixes may have stratum ordering effects in that one group of affixes may, by virtue of their phonological requirements, occur within another set of affixes; however, compounding is not like this in that phonologically it consists in the adjunction of stems or words which do not in any sense phonologically select for one another. So, a compound's syntactic representation is interpreted phonologically in one way or another but there can be no stratum ordering of this process insofar as there is no syntactic stratum ordering and furthermore there is no phonological motivation, such as selection, for imposing an ordering. However, it has been claimed, in particular by Kiparsky (1983a) and Mohanan (1982) that word-formation should be restricted by LPM in precisely this way, and it is to these

claims that we turn directly.

4.2 Against LPM as a Theory of Word-Formation

In this section I discuss a number of cases put forward by Kiparsky and Mohanan which would appear to provide empirical evidence that stratum ordering must be allowed to constrain word formation quite strongly, and that it is consequently not merely a theory of phonological well-formedness as suggested in the last section.

All of the examples that I discuss here have the following characteristic: if the LPM analysis of them is correct and there is therefore evidence that stratum ordering constrains word formation, then there will be a serious problem for my theory in that I will need to also posit some sort of level ordered word-formation device in addition to the checks on syntactic well-formedness which I have argued take place in the syntax. This would make my approach appear baroque and at best a notational variant of LPM. What I hope to show, then is that the examples put forward by the researchers in LPM are either questionable empirically, or explainable by other independently necessary principles.

This section, then, will be a potpourri of examples from the LPM literature and counterarguments to those examples. In particular, I shall

discuss morphological blocking, the inflection of compounds, Malayalam subcompounding and cocompounding and finally one of the more intriguing problems introduced by LPM, namely the treatment, in English, of regular inflections inside compounds (c.f. Kiparsky, 1983a; Mohanan, 1982; Halle and Mohanan, 1985; and against the LPM approach, see Thomas-Flinders, 1983.) The reason for discussing these examples is that they represent the strongest arguments of which I am aware to the effect that combination of morphemes is stratum ordered in and of itself. Again, compounding figures prominently in this discussion insofar as there is no sense in which the constituents of a compound may be said to phonologically select for one another (i.e., they are not specified as phonological affixes.) Rather, compounds on my account have their syntactic properties derived in the syntax, and are merely interpreted phonologically in the phonology; their formation thus cannot be stratum ordered.

4.2.1 Morphological Blocking and Stratum Ordering.

Blocking (Aronoff, 1976) is a common process in morphology (and, I shall suggest in the final chapter, elsewhere too) whereby a more specific word formation process blocks the application of a more general word formation process. To give an example, the irregular plural <u>feet</u> in English blocks the regular plural *<u>foots</u>.

Kiparsky (1983) claims that "an absolute constraint....seems to be that

a blocking process can only be located at these same level (=stratum) or at an earlier level than the process it blocks." The feet/foots example is already illustrative of this point since, according to LPM, an irregular plural such as feet would be formed at a stratum (Stratum I) earlier than the stratum at which the regular plurals are formed (either Stratum III or IV, depending upon the version of LPM.) So, feet would be formed at Stratum I, and at the stratum at which regular plurals are formed, its existence would block the application of the regular plural rule.

Examples such as these are actually not problematic for my approach, coupled, again, with the interpretation of LPM as a theory of phonological well-formedness. To see this, consider a the syntactic representation of the two nouns feet and dogs. There is presumably no reason to suppose that there is a difference in the syntactic representation of the two merely because there is an obvious difference in the phonological representation, so they would be represented roughly as follows[4]:

---

4. A good example illustrating the independence of syntax and phonology in inflectional morphology was pointed out to me by Ken Hale (and see also Jeanne, Pranka and Hale (1984)). In Hopi, coordinated clauses allow for optional deletion of the second verb stem, thus stranding the affixes (OBV=obviative):

'ima yi'ti -k-q pii' itam tiwat [e]-ya-ni
(you:nonsg) run:pl-k-OBV then we also pl fut
"You run and then we will also"

What has been deleted here is the verb stem for run, stranding, in particular, the plural suffix -ya. However, the full form of the verb,

(10) [[FOOT] PL] [[DOG] PL]

Now, the affix PL will presumably have as part of its phonological entry, the fact that with certain stems it has an irregular spell out. A partial entry might look as follows:

(11)

PL' = ⟨PL$_{\langle N,O\rangle}$'
{-s,
Umlaut with _foot_, _man_....}⟩

where Umlaut is some set of appropriately defined phonological rules[5]. Now, if we list such phonological rules as Umlaut as applying at Stratum I and regular phonological affixation processes such as the affixation of _-s_ at Stratum III (or IV), we will thus allow blocking to apply; the Elsewhere affix _-s_ will be blocked by the presence of a phonological form such as _feet_ already spelling out the plural form of _foot_[6]. If the alternative

---

which would be _yɨ'tɨ-k-ni_ has the plural form of the verb stem, which is suppletive for _wari_ 'run:sg'+_ya_. Clearly, then, the deletion operation takes place over a level of representation where the verb is represented separately from the plural affix, irrespective of the eventual phonological form of the verb.

5. Though it might be more to the point to say that plurals like _feet_ are suppletive for _foot+s_; see Section 4.2.2.4.

6. Not that such a stratum-ordered approach would be necessary for these particular examples; the very existence of the more particular rule of Umlaut would block the affixation of _-s_ anyway, according to the Elsewhere Condition. I know of little evidence that the irregular plural rules of English need apply at any particular stratum in the lexical phonology, with the possible exception of the facts concerning plurals within compounds to

ordering were picked--i.e., Umlaut at Stratum III and regular affixation at Stratum I, the blocking effect would go in the opposite direction and Umlaut would always be bled.

The difficulty would come where two separate affixes are involved which mark the same information. In particular, if one of the affixes is attached at Stratum n and the other at Stratum n+m, m>0, the affix added at Stratum n should, according to LPM, be able to block the affix at Stratum n+m. Blocking should not occur in the other direction, however. If this kind of distribution of data were to show up, there would be a problem for my theory since, in any event, the two affixes should have syntactic representations which have no Stratum Ordering specifications on them; blocking should thus in principle be possible in both directions. An example of this kind of situation would be the two negative affixes _in-_ and _un-_ in English. In LPM where the word formation rules affixing _in-_ and _un-_ occur at different strata--_in-_ at Stratum I and _un-_ at Stratum II--one could argue that affixation of _in-_ blocks the affixation of _un-_. So we do not get *_unpossible_ because there is already _impossible_. Supposedly, however, the blocking does not go the other way; that is, there is no sense in which _un_ can ever be said to block the affixation of _in-_.

---

be discussed below. So there may be no reason to assume that there is a stratum ordering distinction between Umlaut and the regular affixation. This example was merely given as an illustrative example of how this would all work.

Again, in my approach such facts might be difficult to account for. The syntactic representation of both words would be:

(12) [IN [POSSIBLE]] [UN [POSSIBLE]]

In the PF component these two forms would be represented with the appropriate phonological representations at the appropriate strata. Now, of course, IN could well block the affixation of UN in the syntactic representation; presumably POSSIBLE is simple marked to take IN and this fact blocks the affixation of UN. The question is why the blocking could not go the other way; why could a form not be marked to take UN and thus block the affixation of IN? In fact, cases which look remarkably like this do occur:

(13)

unable	*inable
unequal	*inequal
ungrammatical	*ingrammatical

.....

Presumably all of the base adjectives in these forms must be available at Stratum I since all of them must take -ity[7]. So why do they not take in-?

---

7. I assume that inability and inequality must simply be listed in these forms. As suggested by Archangeli (1984a), the distribution of negative prefixes in the able/equal paradigm could be explained by stratum ordering insofar as -ity attaches to the adjective marked with the negative prefix; the latter could therefore not be un- because of stratum ordering. Of course, this explanation evaporates as soon as we are forced to account for the much larger class of Bracketing Paradoxes.

The answer would seem to be that they do not take in- because they are marked to take un-. Yet if this is correct, then we have a case of blocking counter to the direction of the ordering of the strata. This seems perfectly natural in a theory where word formation (i.e., the composition of the morpho-syntactic representation of words) is not Stratum-ordered.

4.2.2 On the Stratum Ordering of Compounds.

4.2.2.1 Inflection of Verbal Compounds.

Kiparsky (1983) observes that we do not get irregular inflection of verbs which have been derived from nominal compounds. For example, the compound noun grandstand may also be converted into a verb to grandstand As Kiparsky notes, this verb does not receive an irregular past tense inflection despite the fact that it ends in stand which would normally be irregular: *grandstood. versus grandstanded. Within (Kiparsky's version of) LPM, this would be accounted for as follows. At Stratum I, verbs are converted into nouns. At Stratum II, compounding occurs, as does noun to verb conversion (a process which is very regular and productive in English.) Finally, at Stratum III regular inflection takes place. This is summarized in the following diagram (adapted from Kiparsky):

(14)

Stratum I:	$stand_V$	----> $stand_N$
Stratum II:	$grand_A$ + $stand_N$	----> $grandstand_N$
	$grandstand_N$	----> $grandstand_V$
Stratum III:	$grandstand_V$	----> grandstanded

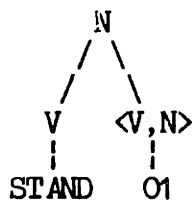
Clearly, since the verb grandstand is formed after the stratum at which irregular verb inflection is available there is no question of irregular inflection of such forms.

Nevertheless, I believe that there is a straightforward account of these facts which does not invoke stratum ordering. The basic assumption which we need--and which I feel is completely justified by the facts--is that irregular verbal inflection, irrespective of the level at which it occurs, must be listed with the particular verbs which take it. The most straightforward justification for this assumption that I know of is that new verbs which are coined (no matter what their source) never to my knowledge receive irregular inflectional morphology. If I make up a verb to zing meaning, say, 'to hit with an aardvark's snout,' it would not be inflected irregularly, with past tense *zang, but would rather receive regular inflection; this is despite the claim that Kiparsky (1983a) makes that verbs ending in the string /ing/ are virtually exceptionallessly inflected strong.

Armed with this assumption, we can now derive the observed pattern.

First of all, we may assume that the noun stand is derived (in the morpho-syntax) from the verb stand by an abstract affix--i.e., one which has no phonological content--which we will call O1. So the noun will have the following representation

(15)



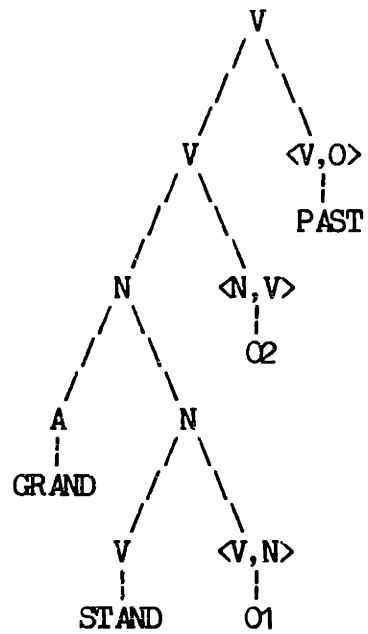
Let us assume that the entry for an affix like O1 will be as follows:

(16) O1'=<O1,0>

The phonological side of the affix O1', namely O1 is neither a suffix nor a prefix (such a designation would, after all, make little sense for something with zero phonological content.)

Assuming that there is another abstract affix, which we will call O2, which converts nouns into verbs, the morpho-syntactic structure of the entire verb form grandstanded would be as follows:

(17)



Given the Mapping Principle (and assuming that PH(PAST) is a suffix, and that there is some principle--such as righthand headedness--to order PH(GRAND) and PH(STAND)), we have the following direct phonological encoding of the above syntactic structure:

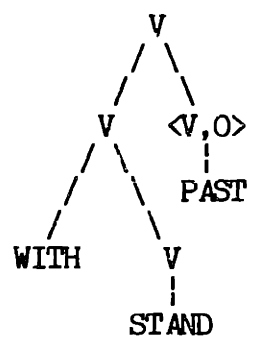
(18) PH([[[[GRAND][[STAND] [O1]]] [O2]][PAST]])=
((((grand)^((stand)*O))*O)^PH(PAST))

The question as to why the phonological spellout of PAST cannot "parasitize" off of the irregular morphology of stand is now answerable: in order for this to happen, PH(PAST) would necessarily have to be phonologically adjacent to stand--i.e., in the technical sense introduced in the first chapter, stand^PH(PAST). Since * is not an associative operator, this will not be possible. (For a discussion of the mathematical

properties of ^ and * see Chapter 1.) So, it is perfectly possible to come up with a reasonable account of the facts Kiparsky discusses without resorting to the assumption that word formation itself is stratum ordered.

Note, too, how the grandstand case differs from a case like withstand which does have irregularly inflected stand. In this case the syntactic representation is:

(19)



This converts to the following:

(20) ((with^stand)^PH(PAST))

In this case, given the associativity of ^, reassociation is possible:

(21) (with^(stand^PH(PAST)))

Stand^PH(PAST) will then be spelled out as stood.

Of course, under my approach, it ought in principle to be possible for a verb derived from a nominal compound to be listed as having an irregular past tense. I have, in fact, found an example which would appear to be of

that form. The verb to hamstring is apparently derived, at least synchronically, from the noun hamstring; the meaning of the verb is 'to cut the hamstrings' and would thus appear to be derivative from the noun in the same way as to elbow is derivative from the noun elbow. Now, hamstring, the noun, would appear to be a compound, derived from string and ham (referring to the back of the leg or thigh, though I suppose that this probably has the status of a cranberry morpheme for many or most speakers); note that the form has compound stress:

(22)

1 3
hamstring

Interestingly, at least according to the American Heritage Dictionary, the past tense of this verb is hamstrung. It is of course possible under my approach to list this verb as having an irregular past tense, although this is surely a marked option. It is not so clear however, that this could easily be accommodated in LPM[8].

---

8. Support for the analysis of hamstring_V given here comes from the following quotation from Fowler (1926):

> With h[amstring], no doubt of the right form is possible; in to hamstring, -string is not the verb string; we do not string the ham, but do something to the tendon called the hamstring; the verb, that is, is made not from the two words ham and string but from the noun hamstring; it must therefore make hamstringed.

Fowler, who was a prescriptive grammarian, was clearly consciously aware of

Kiparsky also discusses another set of data which, although they have nothing to do with compounds, reduce to the same analyses as given above. Within LPM, as noted before, verbs may be derived from nouns at Stratum II. As with the compound grandstand, such nouns may ultimately be derived from verbs. An example of this is the irregular verb ring, which has the related noun ring[9], from which the verb ring (as in 'ring the castle with troops') is derived. This latter verb is, of course, inflected with regular inflectional morphology. Within LPM this fact is captured in exactly the same way as the previous case with grandstand: since the second verb ring is derived at Stratum II, after the irregular verbal inflection is available, it will follow that ring cannot have irregular inflection.

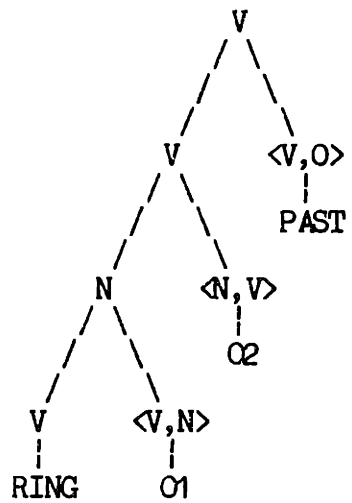
Needless to say, the same approach that I outlined above within my own approach for the verb grandstand would carry over to this example. Ringed, the past tense of the second verb ring, would have a morpho-syntactic representation such as the following:

---

both the status of the verb hamstring and of the generalization encoded in LPM and also derivable frm my approach. Nevertheless, despite Fowler's concerns, and presumably on the analogy with string, hamstring has come to be listed with the irregular past tense; I assume, however, that speakers are no less aware of its derived nature.

9. Though this relationship is admittedly only historical. In any event, the explanation given below for why $ring_V$ derived from $ring_N$, does not irregularly inflect carries over to the case where $ring_N$ is not derived from irregular $ring_V$.

(23)



The direct mapping of this will be:

(24) ((((ring)*O)*O)^PH(PAST))

Again, due to the non associativity of *, this cannot be reanalyzed into a representation where PH(PAST) is adjacent (^) to ring. Hence there is no reason to expect that the second verb ring would be inflected irregularly. Of course, it is still possible for such an entry to merely be exceptionally marked as being so inflected, just as hamstring has to be so marked. I have as yet been unable to find such examples, however.

4.2.2.2 Inflection of Nominal Compounds.

Another set of compounds which Kiparsky discusses are the exocentric or bahuvrihi compounds. Some examples of these follow:

(25)

sabertooth
redcap
skycap
whitecap
whitewall

What makes these compounds exocentric is that neither of the members of the compound, in particular the righthand member, is the semantic head. So sabertooth, unlike milktooth or wisdom tooth is not a tooth, but rather a kind of cat. A whitewall is not a wall, but a kind of tire.

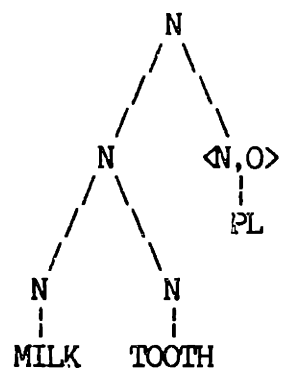
The notable thing about these forms is that they all inflect with regular morphology. As Kiparsky notes:

> ...exocentric (bahuvrihi) compounds are characteristically inflected at level 3 even if their second members are by themselves inflected at level 1, whereas endocentric compounds retain the inflection that their second member has by itself. Consider e.g. milk teeth (endocentric) vs. sabertooths 'sabertooth tigers' (exocentric). As in the verb compound just discussed [i.e., grandstand RS], endocentric compounds are formed at level 2 by combining words, including words derived at level 1 such as teeth. Exocentric compounds however, must on our assumptions be assigned zero derivational suffixes since they otherwise would share the properties of their heads, i.e. be endocentric. But....derivational suffixes cannot be added to derived plurals [a point which Kiparsky notes elsewhere, RS]. Therefore exocentric compounds come out of level 2 with exclusively singular morphology and can receive plural endings only at level 3 where they are adjoined to the whole compound.

Again, however, there is a ready explanation for the facts which does not appeal to stratum ordering. First of all, note that it is perfectly easy to account for the fact that endocentric compounds containing heads which inflect irregularly for plurality also inflect irregularly. In fact,

unlike LPM, it is not necessary to assume that a form like milk teeth is derived separately from milk tooth. The syntactic structure of the plural form of milk tooth would be:

(26)



The Mapping Principle yields the following representation:

(27) (((milk)^(tooth))^PH(PL))

which is convertible to:

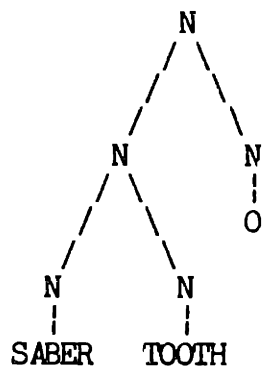
(28) ((milk)^((tooth)^PH(PL)))

So this is just like withstood.

Turning now to the exocentric cases, let us assume that Kiparsky is correct in positing an abstract head for these compounds. In fact, although the argument will not be affected by this decision, it may be possible that this abstract head is a full-fledged noun, which we will call O; after all, saber-tooth is in some sense "short for" sabertooth tiger, so

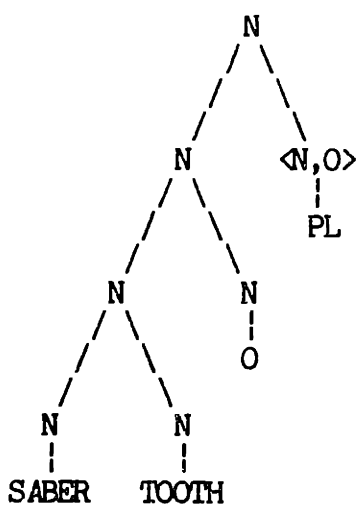
we might assume that the righthand member of the latter construction is still present in the former in the sense that it acts as the syntactic head. The structure for sabertooth would therefore be as follows:

(29)



The structure of the plural form would therefore be:

(30)



Again, by the Mapping Principle, this will yield:

(31) (((saber)^(tooth))*O)^PH(PL)

Assuming, again, that there is no justification for linearly ordering a null phonological entity, it will again follow that the phonological spellout of the plural will not be rebracketable with tooth, due to the non-associativity of *.

So, as for the verb grandstand, we have an explanation for the fact that bahuvrihi compounds do not irregularly inflect even though the righthand member may well do so. Again, however, exceptions should in principle be possible: tenderfoot with the secondary plural tenderfeet provides an example of such an exception.

In the next section I turn to a discussion of some examples of compounding from Malayalam, as discussed by Mohanan (1982).

4.2.2.3 Malayalam Subcompounding and Cocompounding

Mohanan (1982) (pp. 35-64) discusses an interesting set of data from Malayalam which would appear to suggest that a large class of word formation processes is stratum ordered. In that language there are two classes of compounds, which Mohanan refers to as subcompounds and cocompounds. Subcompounds are essentially like English noun-noun compounds; according to Mohanan they have the structure 'modifier + head.' Cocompounds, on the other hand, have a coordinate structure, 'stem + stem + ...;' they refer to the conjunction of the items referred to by the

individual stems. Some examples of each class follow:

(32)

Subcompounds:

a. ṯaarāakaaṉṯanmaarə 'Tara's husbands'
(ṯaarā 'a name'; kaaṉṯan 'husband'; maar̄ 'plural')

b. kaaṭṭəmarām 'forest tree'
(kaaṭ 'forest'; marām 'tree')

Cocompounds:

a. acchanammamaarə 'parents'
(acchan 'father'; amma 'mother'; maar̄ 'pl.')

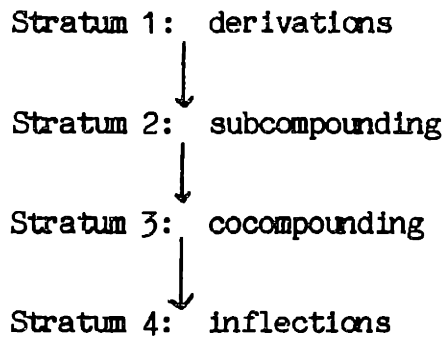
b. yakṣakinnarāgandharwaaḍikaḷə
'Yaksa's, Kinnara's, Gandlanarwa's, etc.'

(yakṣan; kinnaran; gandharwan; aaḍi 'etc.'
kaḷ 'pl.'

Mohanan observes that there are two phonological rules, which distinguish cocompounds from subcompounds. One is gemination; it occurs in subcompounds and not in cocompounds. The other is stress and tone assignment, which apparently occurs at the stratum of cocompounding (see the diagram below), and nowhere else In addition, there are two other rules which occur in both subcompounding and cocompounding, though not derivational or inflectional morphology. These are nasal deletion and vowel sandhi. We return to some of these rules momentarily.

Mohanan suggests that the following be taken to be the structure of the stratum ordered lexicon for Malayalam:

(33)



With respect to compounding, this makes the prediction that subcompounds can occur within cocompounds, which is indeed the case:

(34)

maatrə sneehapatniwidweesam 'mother love and wife hatred'

(maatr 'mother'; sneeham 'love';
patni 'wife'; widweesam 'hatred')

However, it is also true, as Mohanan points out, that cocompounds may occur within subcompounds. It is in fact possible to have a whole sequence of nestings of subcompounds within cocompounds within subcompounds, as the following shows:

(35)

[<{<[maatrə][sneeha]><[patni][widweesa]>}[wikaaraŋ]>ŋalə]
mother love wife hatred emotion plural

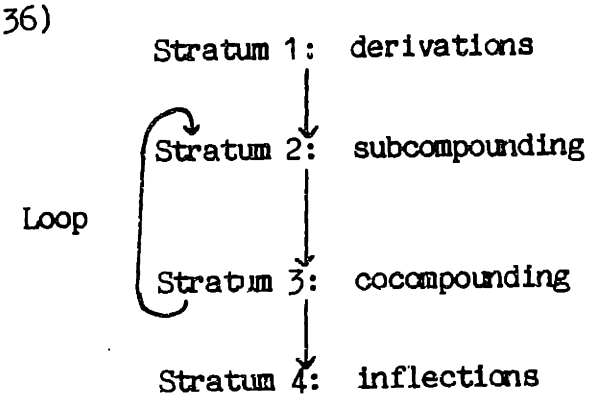
'the emotions of mother love and wife hatred'

(where '◇' delimits subcompounds and '{}' delimits cocompounds

To account for this, Mohanan introduces into the theory of Lexical

Phonology the device of the Loop, which is illlustrated in the diagram below:

(36)



Thus, even though the ordering of the strata in the basic model of Malayalam morphology does not allow cocompounds within subcompounds, the Loop now allows for this possibility since it is possible in the revised model to do an operation of cocompounding at Stratum 3, and then return, via the loop to the second stratum in order to do an operation of subcompounding.

Nevertheless, Mohanan argues, the ordering must still be assumed to be as in A below, and not as in B:

(37)



The argument for this has to do with the behavior of stress and tone in Malayalam, to which we now turn.

The pitch contour for Malayalam is of the form LH, where the L tone is anchored to the primary stressed vowel. Mohanan gives the rules for stress and tone assignment as follows (p. 56):

(38)

Foot Construction

a. Construct feet on all long vowels
b. Mark the first syllable as extrametrical if it has no foot and is followed by a foot.
c. Construct feet on initial and final vowels.

Word Tree

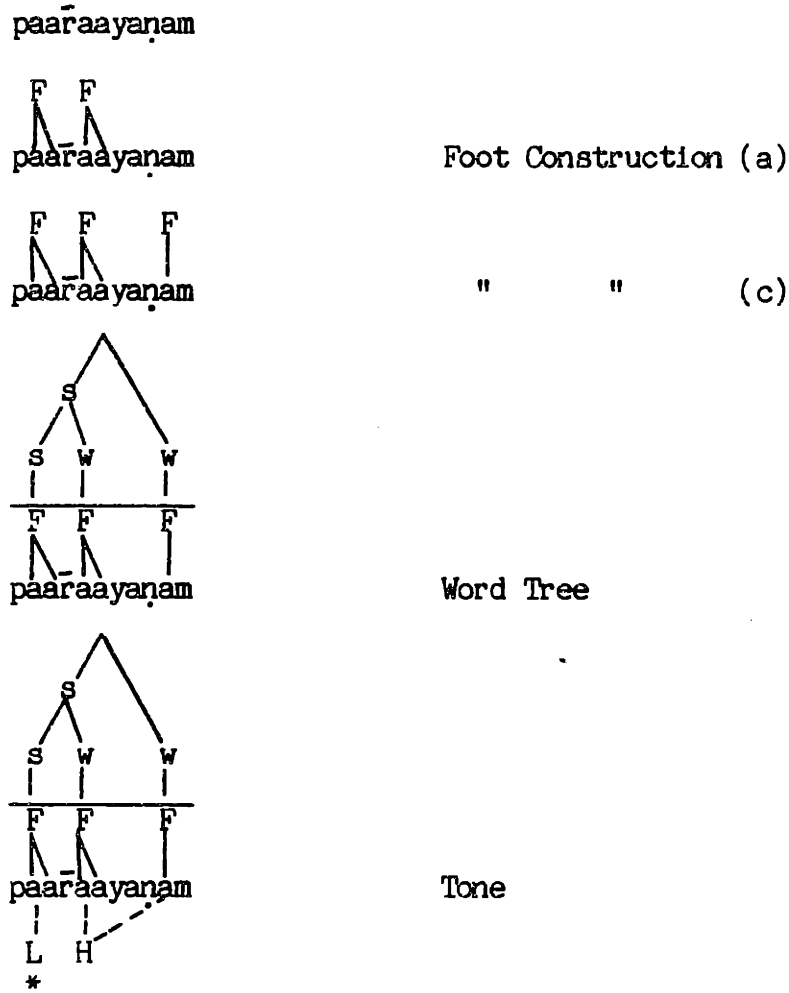
Construct a left branching tree on feet.

Tone

H spreads to all [of the rimes dominated by the: RS] tree.

A sample derivation is given below:

(39)



Interestingly, while subcompounds act as a unit for the purposes of stress/tone assignment, cocompounds do not, as illustrated by the following examples:

(40)

a. Subcompound

[[[táarāà][kàantan]]maarə
L H H H

a name husband pl.
'Tara's husbands'

b. Cocompound

[[[ácchan][ámma]]maar̄ə
L H L H
father mother pl.
'parents'

So, in the subcompound there is only one primary stress, as indicated by the presence of the single low tone, whereas in the cocompound there are two primary stresses.

These facts, plus the fact, which may be observed in the examples above, that inflectional affixes are apparently outside the domain of stress and tone, lead Mohanan to place the rules of stress and tone assignment at the stratum of cocompounding in the organization of the lexicon. So, a subcompound like t̲áar̄àakàaṇt̲anmaar̄ə is formed at Stratum 2, and is then run through Stratum 3 (the cocompounding stratum), where it is assigned a unique primary stress. A cocompound such as ácchanámmammaar̄ə, on the other hand, receives as many primary stresses as there are cocompounded stems. This happens because the stems (acchan and amma) enter the cocompounding stratum, are run through the stress and tone rules, and are then compounded together into a cocompound. Finally, at Stratum 4, inflectional affixes are attached, which are therefore outside the domain of stress and tone assignment.

Note that the facts can only be accounted for under Mohanan's model of

LPM under the assumption that the ordering (irrespective of the Loop) is as in (37A) above and not as in (37B); if the latter were true, then individual stems would be assigned stress at the cocompound stratum (which in the B model precedes the subcompound level) before they are subcompounded together. This would predict, counterfactually, that subcompounds in Malayalam should exhibit more than one primary stress.

We have now introduced the model which Mohanan proposes within LPM for accounting for the facts of compounding for Malayalam and it is time to look more closely at this model. A fair criticism of LPM has to do with the status of the Loop. The problem is that the Loop has a rather questionable status in a theory which would seem at first sight to be making the rather strong (hence interesting) claim that processes of word formation are organized into strata which are themselves strictly ordered among themselves. The introduction of the Loop between two levels is a denial of the second part of this claim, namely that the strata are ordered; that is, phonological processes may well be associated with particular kinds of word formation, but the introduction of the Loop is tantamount to saying that there is no ordering among the particular strata of word formation between which the Loop has been introduced.

However, the Loop might still be an interesting device, if it could be shown that there was evidence from phonology that it existed. That is, if we could show that the phonological form of some output of the lexicon

exhibited evidence that it had visited Stratum m, moved on to Stratum n, and then revisited Stratum m and n again via the Loop, we would have positive evidence that the Loop was something more than a way of encoding the fact that, after all, word formation per se is not constrained by Stratum ordering. I know of no such convincing evidence and the Malayalam data, which are the only data which I know to bear on this question, seem to indicate rather that there is something wrong with the idea that the Loop will show up in the phonological output.

The Malayalam example is one example which Mohanan presents as being in support of his model. Consider the following compound (where, again, '◇' delimits subcompounds and '{}' delimits cocompounds):

(41)

<{[jáa̲ti][má̲ta]}[wi̲dwéesam]>
L H L H L H

caste religion hatred
'hatred of caste and religion'

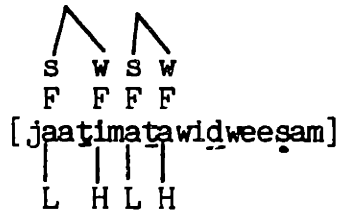
As Mohanan notes, unlike ordinary subcompounds, the interesting thing about this example is that the subcompounding of jaatimatam and widweesam allows for a primary stress on the second member, as evidenced by the L tone on widweesam. Mohanan gives the following derivation for this form (p. 60, = his (60)):

(42)

Derivation	[jaaṭi] [maṭam]	affixation
Subcompounding	——— ———	compounding
Cocompounding	F F F F [jaaṭi] [maṭam] L H L H	Stress and tone
	F F F F [[jaaṭi] [maṭam]] L H L H	compounding
Subcompounding	F F F F [[jaaṭimaṭam] [widweesam]] L H L H	compounding
	F F F F [[jaaṭimaṭa] [widweesam]] L H L H	nasal deletion
Cocompounding	F F F F F F [jaaṭimaṭawidweesam] L H L H L H	Stress and Tone
	———————	Compounding

However, a closer look at this derivation reveals that it cannot be correct according to the principles of Lexical Phonology. The Opacity Principle (Mohanan, (1982)), which has also been termed Bracketing Erasure by Kiparsky (1983c), following Pesetsky (1979), guarantees that at the point at which the second application of the stress and tone rules occurs--i.e., in the second pass through the cocompounding cycle--the internal structure of the compound, which is derived at the subcompounding cycle, is invisible to any rules of phonology and morphology. In other words, the structure is as follows:

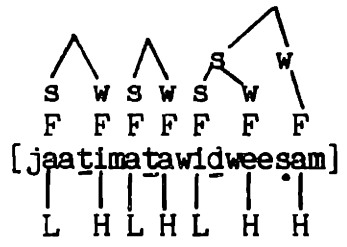
(43)



It is less than clear how the stress rules are supposed to apply to a structure like this. If we make use of the idea discussed by Kiparsky (1983a) that a structure such as (43) counts as underived, and hence cannot be affected by structure changing rules, then we will certainly derive the fact that stress and tone rules applying on the second pass through the cocompounding stratum cannot obliterate or change the metrical and autosegmental structure already present; this much is certainly in keeping with the facts. And we could presumably derive from this that stress and tone rules can only apply to the righthand part of the word over which no

metrical structure is already built. But what we cannot do is to get them to apply so as to derive the correct output. For how can we define the initial syllable appropriately for the application of parts b. and c. of the Foot construction rules? The syllable wid is not the first syllable in the construction in (43); only if there were a left bracket preceding it could it be so defined, yet all of the internal brackets have been deleted in accordance with the Opacity Principle. What is worse is that even if we could define wid as initial, it still could not count as extrametrical since it is not peripheral in the construction (see Hayes (1979); Harris (1983); among many others with respect to the peripherality condition on extrametricality). If it is not extrametrical, however, the stress rules will create a foot over it and we will get the following output:

(44)



This is, of course, wrong.

How then could we derived the correct results within Mohanan's model? In fact, the correct results are derived if we assume that widweesam receives its stress and tone on its own, independently of its role in the compound. According to the stress rules, this word will receive exactly the same tone assignment independently as it apparently does in the

compound. What seems to have happened, then, is that we have looped around through the subcompounding stratum and cocompounding stratum a second time so as to simply mechanically allow for the compound's derivation, and yet there is no phonological evidence to show for the supposed fact that the full compound has made a pass through the stress and tone assignment rules on the cocompounding stratum[10].

Another way around the problem would be to use the following approach. Suppose that the stress rules are given as follows, abstracting away from the tone rules which we take to be the same as in Mohanan's account (with the difference that H spreads to all DTEs):

(45)
Branching rimes are dominant
(1) Foot Construction:

(i) Construct a foot on the final rime.
(ii) Group the remaining rimes into left dominant unbounded feet.

(2) Prestress Destressing: Delete a nonbranching foot dominating a non-branching rime when it precedes a branching rime.

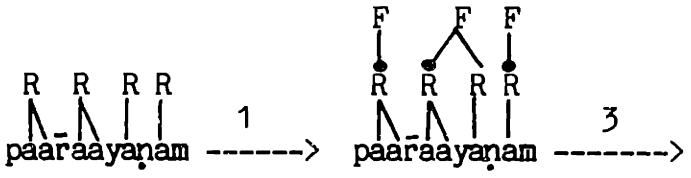
(3) Build a left-dominant word tree on the feet.

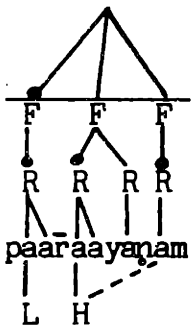
---

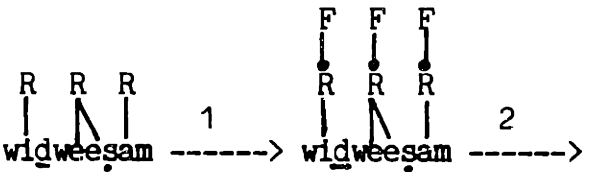
10. Of course, were metrical structure to be derived on widweesam before its compounding with jatimatam, the form would still count as underived on the final pass through the cocompounding cycle and the stress and tone rules could simply not apply since they would otherwise change structure. Nevertheless, this is as good as saying that there is no evidence that they are even given a chance to apply.

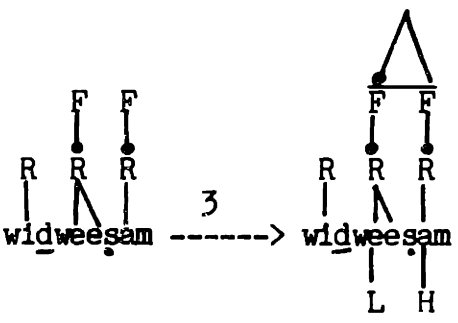
For example:

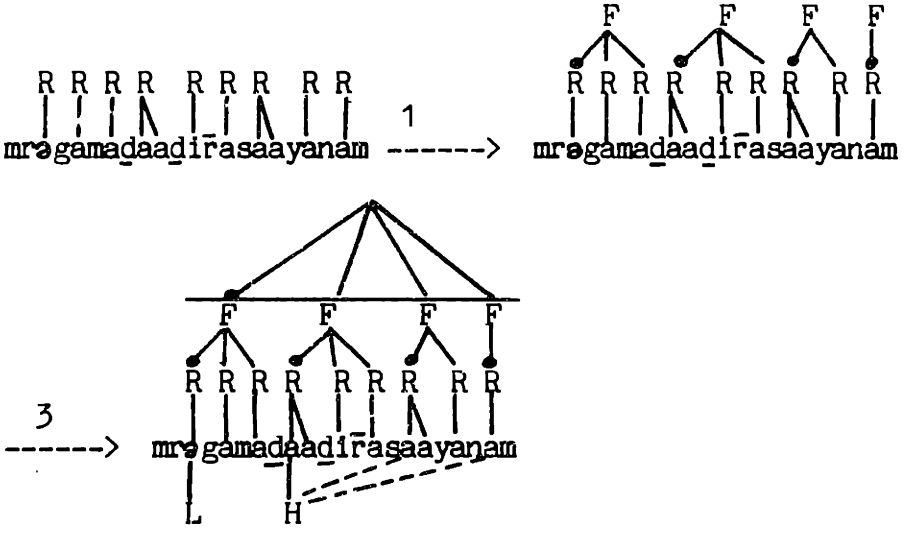
(46)





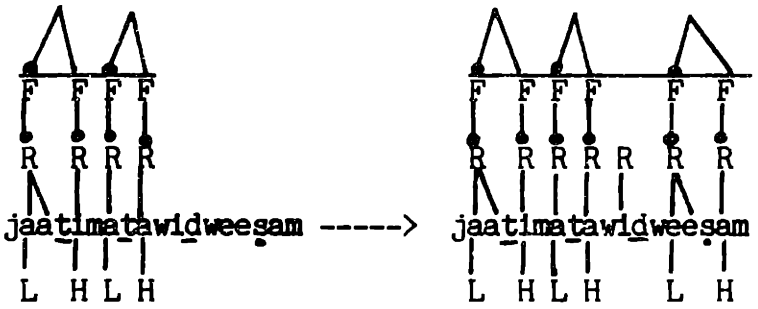






This account is able to derive the facts, and does not make use of extrametricality. Therefore, the objections to Mohanan's system with respect to extrametricality would not carry over to this account. In particular, for jaatimataawidweesam we would simply build metrical structure over that part of the word which had not been previously assigned structure:

(47)



Of course, this would still be indistinguishable from the situation where

jaaṭimaṭam and widweeṣam both had their stress patterns assigned previously to being compounded together.

Still, there is another problem for Mohanan's LPM analysis of Malayalam. Given that separate words must anyway pass through the cocompounding stratum to be assigned stress, and given also the existence of the Loop, what principle is to prevent the following derivation?

(48)

Cocompounding:	Stress & Tone	[maṭa] [widweeṣam] L H L H
Subcompounding:		[[maṭa] [widweeṣam]] L H L H
Cocompounding:		[maṭawidweeṣam] L H L H

So, maṭam and widweeṣam need to be able to pass through the cocompounding stratum in any event so as to be assigned stress as separate words. But given that, what is to stop us returning to the subcompounding stratum and compounding the two stressed words together? This derivation, needless to say, yields incorrect results, but it is not obvious how we rule it out in principle.

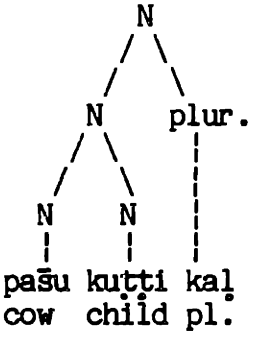
Turning now away from the LPM account of Malayalam subcompounding and cocompounding, how can we provide an analysis of these data which is both able to account for what LPM is able to acccunt for, and does not suffer

from conceptually strange formulations such as the Loop or from technical difficulties such as Mohanan's account of stress and tone within compounds?

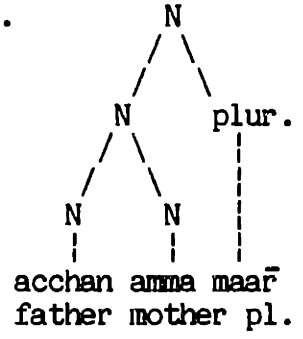
Our starting point will be what I shall posit as the syntactic structures of the two compound types. With respect to this question Mohanan has the following to say (p. 64):

> The assumption that lay behind our analysis of subcompounds and cocompounds was that they are generated by rules (66)a and b respectively:
>
> (66) a. N ---> N N (subcompounds)
>
> b. N ---> N* (cocompounds)[11]
>
> There are no node labels that distinguish between the two kinds of compounds, such as N^{sub} and N^{co}. Thus, the subcompound pāsukkuṭṭikaḷə 'calves', and the cocompound acchanammamaārə 'parents' will have identical configurational structures:

(67) a.



b.



---

11. Presumably Mohanan really means N ---> N N N* here since there must actually be at least two members of a cocompound for it to count as a cocompound; this is a different rule from one like VP ---> V NP PP* where it is really intended that there may be any number of PPs including none.

> The distinction between subcompounds and cocompounds is made in terms of the stratum at which the stems are put together. (66a) has the subcompounding stratum as its domain, and (66b), the cocompounding stratum.

Mohanan goes on to discuss how subcompounds and cocompounds are distinguished semantically by having the appropriate rules for semantic interpretation apply at the particular stratum.

Nevertheless, it is by no means obvious that the assumption which Mohanan makes is justified. Suppose in fact that there is a structural difference between subcompounds and cocompounds; then we could argue that the different phonological behavior of these two classes of compounds is derived from the fact that the different syntactic structures are encoded differently phonologically. This is precisely what I shall now argue is going on.

Consider English constructions of the following form:

(49)
father and-son team
meat-and-potato eater
dog-and-cat fancier
foot-and-mouth disease
knife-fork-and-spoon cleaner

In all of these cases a set of conjoined N's occurs within a compound. The interesting thing to note about the left hand parts of these compounds is that, as far as I can tell they all have fairly level stress among the conjuncts, and certainly do not have the primary stress of the conjunct falling on the first member:

(50)

```
 1        1        3
foot-and-mouth disease

  1    1          1     3
knife-fork-and-spoon cleaner
```

This fact suggests that such English coordinate constructions are not Ns but rather N's, a conclusion further suggested by the ungrammaticality of the following, with the reading that -s has scope over the collective:

(51)

*father-and-sons
*dog-and-cats
*foot-and-mouths
*knife-fork-and-spoons

Inflectional affixes in English presumably only attach to X^0 categories, and this would thus rule out the constructions in (51) under the assumption that they are N's[12].

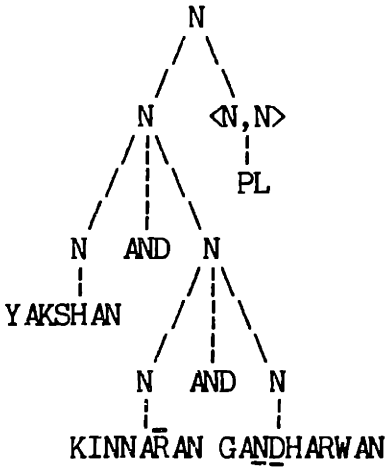
Now, suppose that from a syntactic point of view the Malayalam cocompounds are identical to the English conjuncts in every respect except that the resulting complex is an N rather than an N'. Assuming binary branching (ignoring, of course the conjunction)--and I can see no argument against it--the structure for yakṣakinnarāgandharwanmaarə 'Yaksha's,

---

12. Though Ken Hale has pointed out to me the example ball-and-chain (= 'responsibility'), which has the plural ball-and-chains. This may well be an N^0.

Kinnara's, and Gandharwa's' would be[13]:

(52)



The most obvious phonological difference between Malayalam and English, then is going to turn out to be that in the former language, the syntactic conjunction is not phonologically spelled out in these constructions.

Now, this structure for Malayalam cocompounds certainly has the effect of immediately giving the right semantic interpretation. According to Mohanan, these compounds consistently have the interpretation 'X and Y and Z and...,' something which follows automatically from assuming the syntactic structure in (52) above[14]. Thus we would not need to assume

---

13. I am ignoring for the present discussion the question of how cocompounds might be represented in terms of Goodall's (1984) theory of coordinate structures.

14. Whitney (1889) observes (p. 485) that the same is true most of the time of Sanskrit dvandva compounds, which are analogous to Malayalam

special semantic rules which apply at the level at which cocompounds are formed[15].

Suppose now that we make the following four assumptions: (1) Stress and tone apply at the word level in Malayalam. (See Kiparsky (1983b) on the notion of word level, and see also later in this chapter.) In particular, a phonological word in Malayalam is anything that has a stress pattern assigned by the rules of stress and tone assignment. (2) Cocompounds (phonologically) are conjunctions of (fully-stressed) words. Hence stress and tone rules must have applied before cocompounding takes place. (3) Subcompounds have no such requirements. In fact they must compound (unstressed) stems. (4) Inflectional affixes only attach to fully formed words. They are themselves thus exempt from stress.

Before we see how these assumptions help us, let us justify them. Assumption (1) is fairly straightforward; stress and tone in Malayalam, unlike their counterparts in English, are not sensitive to the morphological structure of words. Hence they are not cyclic rules and can be analyzed as applying as late as the word level. They cannot be later, however, since words are stressed according to the phonological structure of the word and not of the phrase in which the word is embedded.

---

cocompounds.

15. The latter phrase 'at the level at which cocompounds are formed' would, of course, be nonsensical in the model I am proposing anyway.

Assumption (2) requires a little more justification. Booij (1983) argues that coordination reduction across words is a rule which refers to prosodic structure. In particular, in the following Dutch examples involve reduction of a phonological word (on which see Selkirk (1980) and also below):

(53)

a. ontwikkelingen en verwikkelingen ===>
'developments and complications'

ont- en verwikkelingen

b. onnationaal of zelfs antinationaal ===>
'unnational or even antinational'

on- of zelfs antinationaal

c. zwangerschap en moederschap ===>
'pregnancy and motherhood'

zwanger- en moederschap

What also seems clear in fact, is that the elements which are left behind apparently must at least be phonological words. Note the following English cases (see also the short discussion of this topic in Chapter 3):

(54) anti- and pro-abortion, father and motherless, pro- and counter-analysis

In each case, the "stranded" form must be a phonological word in that it must be able to bear stress and furthermore "look like" a word phonologically. Presumably this would account for the fact that forms such as *grammàticál- and ungrammaticality are always bad (contrast with this

the far better ?grammatical and ungrammaticalness.) Say, in fact, that there is a universal condition on coordinated structures:

(55) Coordination Condition.

Coordinated elements must minimally be phonological words.[16]

Assumption (2) will now fall out from this plus the assumption that Malayalam cocompounds are nothing more than conjunctions of (syntactic) N^0.

Assumptions (3) and (4) are fairly straightforward. Assumption (3) is necessary to derive the fact that basic subcompounds apparently act as if they are one word for the purposes of stress and tone. If we assume that subcompounding occurs on stems then we can capture this behavior since the stress and tone rules will therefore treat the string of phonological stems as one unit in forming a word. (4) merely states the generalization that in Malayalam inflectional affixes also must attach to phonological words. Note that (3) and (4) may well have to be language specific (learned) properties of Malayalam. On the other hand, with respect to (4) it seems to be common for inflectional affixes to attach to phonological words; this is apparently true in English and also in Kannada (Aronoff and Sridhar, 1983). Further, note that there is a general prohibition on inflection within compounds in Malayalam. This is, of course, accounted for in the

---

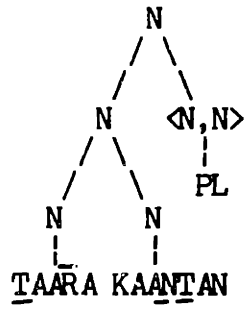
16. Donca Steriade has pointed out to me that this condition may not hold of Classical Sanskrit dvandvas, which have one stress per compound.

LPM model by stratum ordering. However, we might also imagine that there is a semantic constraint which requires that nominals within compounds be unmarked for number. As we shall also see in the next section, this constraint plays a large part in English despite the fact that morphological plurals do occur within English compounds.

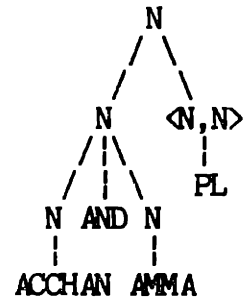
To see how this all works, let us first discuss the following simple examples: the subcompound taaraakaantanmaara 'Tara's husbands,' and the cocompound acchanammamaara 'parents' (lit. 'father(and)mother+pl'). These will have the following syntactic structures:

(56)

a.



b.



Given the Mapping Principle, these will map directly to:

(57)

a. ((ṯaāra^kaaṉṯan)^maārə)

b. (& (PHI(AND)*acchan)
(PHI(AND)* amma)
(acchan^amma))^maārə[17]

and since PHI(AND)=0, (57b) essentially reduces to:

(58) (acchan^amma)^maārə

Assuming that there is nothing forcing a different phonological bracketing from the one given we may assume that the Default Phonological Bracketing Principle applies and that the phonological bracketing is therefore congruent to the syntactic bracketing.

In the case of the subcompound, we assume that the phonological entries of ṮAARA 'Tara' and KAAṈṮAN 'husband' are listed as stems[18]. Hence we merely concatenate the two stems together. However, in order to attach the plural affix -maarə to the compound, we require (assumption (4)) that the compound be converted from a stem to a phonological word, which will entail that it be assigned stress. Therefore we will end up with the tone pattern:

---

17. The conjunction & here stands for the assertion that all of the conditions mentioned must hold.

18. See, again, the final section of this chapter for a discussion of the stem/word theory I am assuming here.

(59) táar̄àkàan̲tanmaar̄ə
L H H H

The same requirement would be invoked, of course, if we were merely to insert the word directly, without an inflectional affix, into a phonological phrase. Such is the case, for instance, in mat̲awid̲weeṣam 'religion hatred,' where no inflectional affix is added, but it still must be the case that the compound be a phonological word. Note, also, that the (Malayalam specific) statement that only stems may be subcompounded (Assumption 3) removes the possibility of a derivation, such as that given in example (48), where it was shown that Mohanan's model would allow doubly stressed subcompounds. Since stems are stressless by definition, it will follow that there can only be one primary stress per compound.

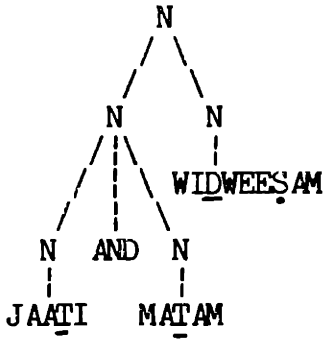
Turning now to the cocompound, recall that the Coordination Condition will require that the phonological spellouts of the conjuncts be phonological words in order for the whole construction to count as a correct phonological representation for the syntactic representation. Thus, in acchanammamaar̄ə we will require stress assignment on both acchan and amma, before they can be conjoined together. This will predict that the compound will have the stress pattern given below, which is, of course, correct:

(60) ácchànámmàmaar̄ə
L H L H

In particular, -maarə will be satisfied insofar as it will be attaching to a stressed item, hence to a word.

Let us now turn to the compound jaatimatawidweesam 'hatred of caste and religion.' I assume that the structure of this is as follows:

(61)

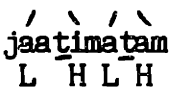


This will ultimately translate to:

(62) ((jaati^matam)^widweesam)

The cocompound jaatimatam will, as in the previous example, have to receive stress on both conjuncts yielding the following tone pattern:

(63)



The following question then comes up: how does the stem widweesam receive stress when it is attached to jaatimatam? Presumably it must do so in order for the whole compound to count as a phonological word (which it must do at some point before it becomes part of a phonological phrase), but

where does this take place? In fact, we can just assume the derivation given in (47), whereby metrical structure is assigned to that part of the word which does not already have metrical structure, namely widweeṣam.

For a more complicated example such as maatrəsneehapatniwidweeṣawikaarāṇṇalə 'the emotions of mother love and wife hatred' I predict the following tonal pattern:

(64)

```
mother love   wife  hatred   emotion  pl
maatrə-sneeha-patni-widweeṣa-wikaarāṇ-ṇalə
 L      H  H  L       H  H    L  H
```

Mohanan does not give the tonal pattern for this word, however.

Thus, we have provided an account of Malayalam compounding without appealing to the idea that word formation--i.e., the syntagmatic combination of morphemes--is itself Stratum ordered. It remains only to note briefly that the rules of Nasal Deletion, Vowel Sandhi and Gemination can be accounted for fairly straightforwardly. The first two rules occur with both subcompounds and cocompounds (the reader is referred to Mohanan's discussion for examples) and I shall suggest that these rules occur automatically whenever the phonological reflexes of two N^{o}'s are concatenated together. Gemination is trickier insofar as it only occurs in subcompounds (and only in Dravidian stems.) However, a morphophonological process occuring only in ordinary compounds--"compound formatives"--is by

no means unprecedented; see Massam (1983), for instance, for a suggestion that lenition in Irish is used as a "compound formative."

In the preceding discussion, I have tacitly assumed the following basic principle governing the application of phonological rules: Phonological rules apply only when they have to. For instance, if a rule is defined as being cyclic then it must be tested for application on every cycle. If a rule is defined as being related to a particular affixation then it must be tested for application whenever that affixation occurs. Finally, if a rule is defined as being a criterion for well-formedness of some phonological constituent--e.g., root, stem, word--then it will be tested for application to a form whenever it is required that that form be interpreted as a member of the particular class of phonological constituent. So, in Malayalam, a sequence S of stems concatenated in a subcompound must undergo the rules of word-level phonology, in particular stress and tone assignment, in order to allow the affixation of an inflectional affix, or to allow the further compounding of these stems into cocompounds. I return to a fuller sketch of such a model of phonology in the last section; however, given that the whole sequence S must be a phonological word, stress rules only need apply to the whole of S and never to subparts of it.

I do not claim in this discussion to have solved all of the intricacies of Malayalam compounding and its phonology. I am certain that there must be much more to be discussed on this topic. What I have done, however, is

to try to give an account which does not simply describe the facts by stipulating the ordering of word formation processes. In many ways the LPM model seems alluring as a theory of word formation and its interaction with phonology in that it manages to account for much data in an apparently neat and simple way. The predictive power of an LPM model for a particular language can be quite powerful too. Nevertheless, it can easily be overlooked that Stratum Ordering merely stipulates the structure of words rather than explaining that structure. Of course, in some cases we might be able to do no better than merely stipulate a particular ordering. This is not always the case, however, and it is in particular, not the case with Malayalam compounding as I have shown. So why should the stress and tone rules interact in the way that they do with the two types of compounding? LPM has an answer which amounts to stipulating the observed phenomenon. It is a relatively unenlightening point, for instance, that the subcompounding and cocompounding strata in Malayalam happen to be ordered in the way that they are and that the rules of stress and tone assignment occur where they do; one could easily imagine another arrangement. However, when we try to account for the facts without positing this machinery, as I have done, we are forced to look more deeply at what might be going on. The story given here for the difference in subcompounds and cocompounds not only appears to be able to account for the facts, but it also links the cocompounding structure in an apparently insightful way to superficially different coordination constructions in other languages.

4.2.2.4 Inflections within English Compounds

We now turn to a discussion of the problem of regular versus irregular inflectional morphology within English nominal compounds.

It has been known for a long time (Jespersen, 1909-49; Marchand, 1969) that the left hand member of a nominal compound in English is generally uninflected, no matter what the interpretation of that member is intended to be with respect to plurality. So consider the following examples:

(65)

dog-lover	*dogs-lover
rat-chaser	*rats-chaser
log-cutter	*logs-cutter
hand-made	*hands-made
finger-bowl	*fingers-bowl
coat-rack	*coats-rack
three-twinkie lunch	*three-twinkies lunch

So a dog-lover is a person who, presumably, loves dogs in general, rather than just one dog. Similarly, a rat-chaser must surely chase more than one rat in order to qualify as a rat-chaser, yet the singular form is apparently used anyway. On the basis of such data one would be tempted to state the following constraint[19]:

(66) The left member of a compound must be uninflected for number.

---

19. See Churma (1983), however for a discussion of the idea that constructions of the form Ns-N are "phonostylistically" bad.

English overtly marks plural number only. Singular is unmarked and is therefore phonologically identical with the uninflected form. So, given constraint (66) it will appear as if English has singular left-members even when those left members are interpreted as plural. This constraint is also presumably at least consistent with the interpretation of members of compounds as having generic reference.

All this would be fine but for the fact that it has been noticed by researchers within the theory of Lexical Phonology that one can apparently find <u>irregularly</u> inflected plurals within compounds. The following are typical examples (the following list adapted from the discussion in Thomas-Flinders, 1983):

(67)

feet-first
mice-infested
teeth-marks
alumni club
menfolk

Notice, again, the contrasting compounds containing regular plurals:

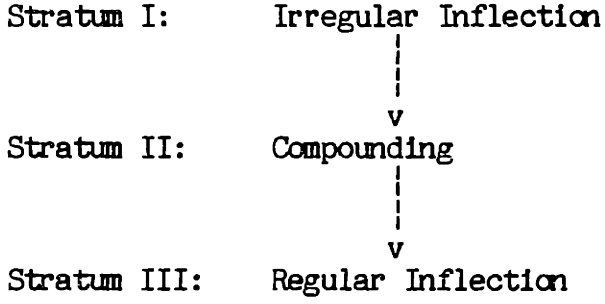
(68)

*heads-first
*rats-infested
*dentures-marks
*students-club
*androids-folk

Supposing that this is the relevant contrast, LPM has a ready explanation: irregular plurals, insofar as they are necessarily idiosyncratic, are

formed at Stratum I; compounding takes place at a later stratum, say Stratum II, and regular inflection takes place at a still later stratum, say Stratum III. The model then takes on the following form (Kiparsky, 1983a; and see Mohanan 1982; Halle and Mohanan, 1985 for a somewhat different (but for these purposes equivalent) model):

(69)



This model, of course can handle the contrast noted above; it could hardly not do so given that it is essentially an encoding of the facts. In particular, constraints such as (66) now seem unnecessary.

Problems would seem to return, however, when we notice that regular plural forms apparently can occur within compounds This point has been noticed by a number of researchers including Kiparsky (1983a)[20], Selkirk (1982), Thomas-Flinders (1983), Hammond (1984), Gordon (1984) the

---

20. He noticed at least that morphologically regular pluralia tantum--i.e. plural nouns such as alms, odds, which have no corresponding singular form (c.f. the ungrammaticality of *alm, *odd)--do occur within compounds: alms-giving, odds-taking

following list is from Hammond (1984):

(70)
systems analyst
Parks department
human subjects committee
ratings book
numbers racket
parts department
jobs program
reservations desk

Needless to say, this fact would seem to be problematic for the LPM model.

Nevertheless, there would appear to be a way out, which has been suggested by proponents of that model (c.f., Gordon, 1984). It can be argued that the plurals in the left-hand parts of the compounds take on an idiosyncratic or collective interpretation. Thus, a systems analyst is not one who analyzes just any old systems, but must analyze systems of some specific sort--e.g. computer systems, economic systems, etc. A Parks-department is a department not of just any old parks, but rather of some specific set of parks. In numbers racket the reading of numbers is clearly idiosyncratic since it refers to a particular kind of gambling based on picking a number from some specified set of numbers.

Assuming that this is correct--and there is clearly something right about it--we can seemingly solve the problem within LPM by proposing that since the semantic interpretation of the plurals in such cases is idiosyncratic, the collocation of the -s ending and the base noun in fact takes place at Stratum I rather than at Stratum III. Stratum I is the

locus for the more idiosyncratic morphological processes within LPM (see, in particular, Kiparsky (1983a) for a discussion of this), and we therefore apparently have good reason for placing such plural formations, along with pluralia tantum, which must apparently be listed anyway, at Stratum I. But, if we can say this, then we have solved the problem since these plurals should, given the organization of LPM given in (69), be able to occur within compounds.

But there is a technical problem which is introduced by this solution. In principle, since such idiosyncratic plurals are derived at Stratum I they ought to be eligible for Stratum I phonological processes; this will hold as long as we take seriously the idea that, as outlined in LPM, morphological, phonological and semantic processes go hand in hand. Nevertheless, I know of no evidence that this is the case. In fact, all idiosyncratic plurals with regular plural morphology, including pluralia tantum have exactly the same phonological processes occuring within them as perfectly regular run-of-the-mill plural forms such as cats, dogs, and fishes have. The /-S/ assimilates in voicing[21]: odds, (/dz/); pants, (/ts/). Schwa insertion takes place when the /-S/ is affixed to a coronal continuant: (sun)glasses, /səz/. This is not merely a metaphysical point; there are phonological processes which take place at Stratum I which could,

---

21. We can assume, for the sake of argument, that the regular plural suffix is unspecified for voicing, though this is hardly a crucial point here.

in principle, apply to such forms, and yet apparently fail to do so.

To see this, consider the treatment of irregular past tense forms in Halle and Mohanan (1985). We are interested here in forms of the following type (their examples (122)):

(71)

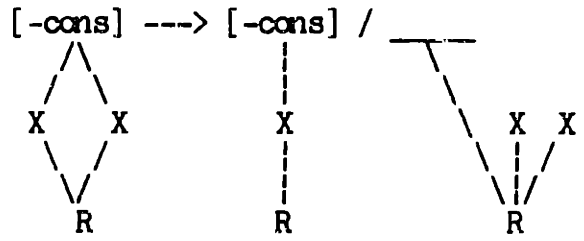
a. bereave--bereft, cleave, creep, deal, dream, feel, keep, kneel, lean, leap, leave, mean, sleep, sweep, weep

b. bite--bit, light, meet

c. lose--lost

According to Halle and Mohanan, these all involve affixation of /t/ at Stratum I. Since this creates a final consonant cluster in these verbs the vowels become subject to Cluster Shortening, which Halle and Mohanan formulate as follows and which occurs at Stratum I:

(72) Cluster Shortening



Where a sequence of two /t/s is created by the affixation, Degemination occurs, but only after the application of Cluster Shortening; Halle and Mohanan in fact place Degemination at the Postlexical Stratum, but this is irrelevant for our purposes. There is also a rule of Voicing Assimilation

which is needed to handle alternations such as bereave--bereft and build--built. A sample derivation is given below:

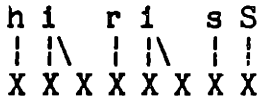
(73)

Stratum I		
Affixation	[[bere:v]t]	[[bi:t]t]
Cluster Shortening	[[berev]t]	[[bit]t]
Stratum II:		
Bracketing Erasure	[berevt]	[bitt]
Postlexical Stratum:		
Voicing Assimilation	[bereft]	_____
Degemination	_____	[bit]

Now, Epenthesis, which accounts for schwa insertion in regular plurals such as laces and also regular past tenses such as rated occurs, according to Halle and Mohanan, at Stratum IV. It cannot apply to a form such as [bitt] at Stratum IV since that is not derived at that Stratum. So one question which comes up immediately is why, if idiosyncratic regular plurals are derived at Stratum I, do we find epenthesis exhibited in them? Why is glasses /glasəz/ and not /glas/? Furthermore why can Cluster Shortening not apply? To show that it does not consider the following example. If a city has a commission in charge of tall buildings I could

well imagine it being termed a highrises commission, in parallel with parks commission and human services commission. Nevertheless, not only has highrises undergone Epenthesis, but it has failed to undergo Cluster Shortening. Yet, under the assumption that such forms are produced at Stratum I, this ought to be possible, since the representation for this form at Stratum I would be:

(74)



Obviously *highris (/hayrIs/) is not a possible output. So, we seem to have strong evidence that plurals inside compounds are not phonologically exceptional as might be expected if they have to be ordered at Stratum I. I do not doubt that there remain ways around this problem which would allow us to maintain that these plurals are derived at Stratum I. Nevertheless, such devices would, it would seem, merely obfuscate the obvious point that there is nothing exceptional about idiosyncratic plurals--even pluralia tantum--from the phonological point of view[22]. Notice that there is a striking contrast between these cases and the examples of irregular past tenses discussed in Halle and Mohanan where there are genuine phonological

---

22. Of course, even if one were to adopt such machinery, the very fact that I can freely form compounds like highrises commission suggests that there is something fishy about assigning such cases to Stratum I, even allowing for the somewhat idiosyncratic or collective interpretation of the plural.

motivations for ordering the irregular affixation at Stratum I.

Given this, we may want to reconsider whether or not LPM really does provide an adequate account for the facts of inflection within compounding in English. Suppose in fact, that we reintroduce constraint (66) in the following form[23]:

(75) The left member of a compound must be unmarked for number, unless the plural is interpreted collectively or idiosyncratically.

This is, needless to say, a rough characterization: this topic has certainly not been investigated sufficiently well so as to allow for a clearer statement of what is going on. Nevertheless, this does seem to provide a fair characterization of at least some of the facts. In particular, the contrast between cases like drink-cabinet versus drinks-cabinet seems do be captured roughly by (75). So, drink-cabinet could be any cabinet in which potable items are stored; drinks-cabinet (as pointed out by Gordon (1984)) seems to refer preferentially to a cabinet for alcoholic drinks, where drinks is interpreted idiosyncratically.

With respect to pluralia tantum, since they have no corresponding singular form, they must merely be idiosyncratically listed as plurals.

---

23. In some ways this is similar at least in spirit to Hammond's (1984) Inflectional Access Constraint.

Therefore, they should regularly occur within compounds; examples like pants-pocket, odds-taking, alms-giver, suggest that they at least can do so. Nevertheless, there is some further idiosyncrasy here: note scissor-blade and trouser-pocket, the latter, in particular, contrasting with pants-pocket (*pant-pocket). This would appear to be evidence that while such forms must be listed as occuring only in plural forms as separate words, the unmarked base must also be a lexical item.

Of course, a principle such as (75) would also predict that irregular plurals should be constrained in the same way, and hence be unusual within compounds in the normal course of events. In fact, this is clearly not an undesirable consequence insofar as although irregular plurals can occur with plural interpretation within compounds, a much larger number of compounds exist with an at least pragmatically plural interpretation of the left-hand member, but where this noun, although it is inflected irregularly, nonetheless occurs in the singular. The following examples are from Thomas-Flinders (1983: p. 130), who points this out:

(76) toothbrush, toothpaste, toothpowder, footgear, footbinding, footpath, footwear, woman chaser, womanhater.

So, a toothbrush is used for brushing more than one tooth, footwear is used on both feet, and a woman chaser is someone who chases after women. Nevertheless, despite the fact that tooth, foot and woman all have irregular plurals, this fact is apparently not made use of in forming these

compounds despite the meaning. In fact, in my experience, speakers usually reject constructions of the form mice-racing on first hearing them and strongly prefer the singular form (mouse-racing): there really does seem to be a strong constraint against using plural forms within compounds except when a collective (parks-department) or idiosyncratic (drinks-cabinet) interpretation is intended[24].

Nevertheless, there still is a problem here insofar as a form like mice-infested is slightly better, if ever so marginally, than rats-infested; at the very least, it seems as though one would have less trouble convincing oneself that the former is well-formed than that the latter is well-formed. This difference is further brought out in an experiment performed by Gordon (1984) which apparently shows that even young children are aware of this contrast. Gordon tested children ranging from 3;2 to 5;10 in age. In the experiment singular, plural and compound forms (the latter all of the form X-eater) were elicited from the children for nouns which were either (i) regularly inflected (ii) irregularly inflected or (iii) pluralia tantum. Even the youngest children of this group knew the regular plural rule as this is something which is learned at a very young age (Berko, 1958). The relevant results were as follows: when

---

24. This constraint even seems to extend to inherently plural words such as people. I remember that the compound people-mover, which I first heard when visiting Disneyland, sounded considerably worse than the alternative person-mover.

children were dealing with regular nouns (or nouns that, through incomplete acquisition, they treated as regular --e.g., feet~feets, mouse~mouses) they practically invariably placed the singular form inside the compound. On the other hand, with irregular plurals, the children invariably placed the plural form within the compound. Now, in the context of the experiment as described by Gordon, in both cases the pragmatic interpretation of the forms within the compound was plural. To be sure, this, plus a certain amount of cueing from the previously elicited plural form must play a part in the explanation for why the children invariably placed the irregular plurals inside compounds. Nevertheless, he correctly observes that this pressure to produce a plural form only works with irregular plurals; in regular plurals it fails, apparently because of some constraint which blocks such forms. The constraint he argues for is the LPM model; in particular, he suggests that much of the structure of LPM must be present to begin with and that therefore children will exhibit behavior consistent with that structure even without much external evidence. In fact, he asks, what external evidence could the child have that irregular plurals within compounds are fine in contrast to regular plurals? As we already have noted, in adult speech neither construction is common. If we assume, however, that children are somehow geared towards placing irregular inflection at an early stratum in the LPM model, and compounding and regular inflection later, we can explain why no learning is apparently necessary.

Still, while this explanation appears to account for the facts, I suspect it is wrong simply because, despite the results of this experiment, plurals of any sort within compounds are not common except in the restricted set of circumstances already described. Nevertheless, we are obligated to provide an alternative story about these facts, especially since the data from the experiment are so striking. The explanation, I believe, is not hard to come up with, if we assume, which I think we must, that irregular plurals are simply listed as the plural variants of nouns already listed elsewhere. Under this interpretation, they are not formed at all by any process, but rather are merely selected as the appropriate form in the translation from syntactic structure to phonological structure. Given that these forms have a somewhat separate existence from the regular singular forms, it is perhaps unsurprising that they could be elicited in compounds; since as Gordon himself suggested, cueing from the previously elicited plural must at least be present in the elicitation of the compounds in the experiment[25] it seems likely that what the children were doing was treating the irregular plurals such as mice and teeth as words more or less divorced from their singular counterparts. They would, of course, still be violating Constraint (75), since such nouns would be

---

25. Otherwise how could we explain why plural forms were nearly exclusively used with irregular nouns within compounds by the children, which is counter to the normal tendency, at least in adults, to use a singular form even when an irregular plural is available?

marked as plural, but this violation would be less obvious than the corresponding violation with regular plurals; in the latter case there is no question that a form like rats is the plural of the noun rat; in the case of mice, however, insofar as it is a separate lexical entry, it could equally well be treated as a noun divorced from mouse, and which just happened to have a plural interpretation. This story, while hardly precise, might well be on the right track towards explaining why, on the one hand, forms like mice-eater are unusual, but, on the other, nevertheless possible. The same story would presumably account for the distinction which adult speakers give to mice-infested and rats-infested.

Why must we assume that mice is a separate entry from mouse? There is evidence for this assertion, although it is, unfortunately, of a somewhat anecdotal nature: Within the past few years an increasing number of high-resolution graphics terminals have been equipped with a gadget which has been termed a mouse, presumably because of its superficial resemblance to said animal. What is the plural of this word? Interestingly enough, a number of people, myself included, have a certain amount of difficulty in making the extension of this idiosyncratic meaning of mouse to the expected plural form mice. Of course, this is the form which is ultimately used; ??mouses seems even more odd than mice. But it is not without some amount of strain that the extension is made. I suppose that there would be no difficulty with a regular plural: were someone to come up with a larger mouse, which might be termed a rat, there would presumably be no difficulty

in forming the plural rats. This anecdotal evidence suggests that there is something in the idea that mouse and mice are separate lexical entries in the sense that they occupy two separate slots in the permanent lexicon more or less as person and people would, and this in turn allows for the style of explanation offered in the preceding paragraph.

In this section, then, I have attempted to offer an alternative to the LPM account of plural nouns within compounds. As with the discussion of Malayalam compounding, I do not claim to have provided a fully adequate account[26]; in particular, I think that a great deal of research needs to be done to determine the exact nature of the Constraint (75), which nevertheless must be on the right track. If my discussion of Gordon's experiment is correct, it seems as if we do not need to assume that the skeleton of a stratum ordered morphology is known by children before they start learning a language. All that they would need to know is something like Constraint (75), which might well reduce to a more basic principle having to do with the interpretation of genericity. Add to that their (acquired) knowledge of listed forms, and we can account fairly straightforwardly for the behavior pattern observed. If all of this is right then, we may have achieved the reduction of yet another class of

---

26. I should say in defense of this, though, that the LPM account is scarcely fully adequate either since it suffers from a number of problems as I have outlined above.

supposed evidence that (syntactic/semantic) morphological operations are stratum ordered. In the next section we turn at long last to an examination of the behavior of the phonology as it is characterized within LPM.

4.3 On the Phonology of Lexical Phonology.

4.3.1 What are the Principles of Lexical Phonology?

I argued in the first section of this chapter that LPM is consistent with the approach to morphology taken here, if we take it to be a theory of phonological well-formedness. In this section we examine how useful a theory of the latter LPM really is. In particular I examine various constraints and principles which have been claimed in the LPM literature to be characteristic of the behavior of the phonology when it applies in the lexicon. In 4.3.2 I shall sketch an alternative phonological theory to LPM.

In this subsection I shall be looking specifically at the following topics: Bracketing Erasure (Opacity Principle) and the arguments for it; Cyclicity, in particular the question as to whether all lexical rules are cyclic and if postlexical rules may be cyclic; and finally whether Structure Preservation can be maintained as a principle of lexical rules.

Each of these principles have been argued to be constraining of the lexical phonology in particular, as opposed to the postlexical phonology. By arguing, as I shall, that Bracketing Erasure is superflous, that Cyclicity is neither true of all lexical rules nor restricted to lexical rules, and finally that Structure Preservation is problematic as a constraint on lexical phonology, we will have reduced the differences between lexical and post-lexical phonology (note the list of such supposed differences given above, section 4.1.1.) Thus, phonology "in the lexicon" is not in principle distinct from other kinds of phonology. I take each of these topics up in turn.

4.3.1.1 Bracketing Erasure.

Mohanan (1982) states the following constraint on the application of lexical rules:

(77) Opacity Principle (Mohanan, 1982)

> The internal structure at one stratum is invisible to the processes at another.

This has been reformulated, in particular by Kiparsky as Bracketing Erasure[27]:

---

27. Pesetsky (1979) formulated Bracketing Erasure as applying at the end of every cycle.

(73) Bracketing Erasure (Kiparsky, 1983c)

> Internal brackets are erased at the end of every level (=stratum).

The consequence of either formulation is obvious any rule of phonology or morphology which applies at Stratum n will be unable to make use of any of the internal structure of words derived at Stratum n-1 or at previous strata. So, for the purposes of Stratum II phonology and morphology in English, a word like organizational will be indistinguishable, as far as its internal morphological structure is concerned, from a word like elephant, which is presumably not morphologically complex.

What I wish to examine is the validity of the arguments that such a principle is in fact necessary; if it is necessary then it seems as if there is a reason to believe that lexical phonology is organized into well-defined strata; if, however, it is not necessary, as I suspect, then the arguments for a specifically lexical manner of phonological rule applications is somewhat weakened insofar as the strata, which are supposedly part of the organization of the phonology, will not have the same constraining effect that they do under the Bracketing Erasure model.

One thing that will be important to bear in mind is how Bracketing Erasure relates to the Strict Cycle (Mascaró, 1976), which shall assume as a valid constraint on the application of cyclic rules. Kiparsky (1983a,b) discusses at length the necessity of assuming that the Strict Cycle does in fact hold, and I shall take the position that this is in fact the case.

The question will arise with at least one of the arguments for Bracketing Erasure as to whether or not it is really not just an argument for the Strict Cycle[28]. But before we examine that point, is there reason to suppose that Strict Cyclicity alone is satisfactory and that Bracketing Erasure may be unnecessary?

Let us examine this question. The two principles do, of course, make different predictions, and it would be useful to compare them. In fact, let us compare the following two theories:

(79)
A. A theory with the Strict Cycle and no Bracketing Erasure.
B. A theory with the Strict Cycle and Bracketing Erasure.

Theory B is basically the theory of Lexical Phonology: both Bracketing Erasure and Strict Cyclicity are supposed to hold. A is a more impoverished theory of the organization of the phonology.

Imagine, then, the following situation: Suppose we have a language which has the following organization:

---

28. This was pointed out to me by Donca Steriade.

(80) Stratum Ordered Phonology of Hypothetical Language A.

Stratum I: Morphology: Affix -a

Stratum II: Morphology: Affix -b
Affix -a' (which is a different affix from -a but phonologically identical with it)
Phonology (rules ordered as given):
Rule 1: a-->q/]__b
Rule 2: a-->r/__b

Consider now the following strings of morphemes and consider how each would in fact be treated under theories A and B:

(81)
(i) xab (monomorphemic)
(ii) xa+b
(iii) x+a+b
(iv) x+a'+b

In the first case both theories would make the same prediction since the form xab is monomorphemic: it would pass through Stratum I, where no morphology would take place. It would then pass on to Stratum II, where, again, no morphology would take place and none of the phonological rules would apply: Rule 1 could not apply anyway since its structural description is not met; Rule 2 could still not apply since it would have to violate Strict Cyclicity to do so, assuming that Strict Cyclicity disallows application to forms that are not derived with respect to the current cycle (of which there is none, in this case, since there is no operation of morphological concatenation.)

Now consider the second case. Again, both theories will make the same prediction. The stem xa will pass through Stratum I, where nothing will take place. At Stratum II, affixation of -b will take place yielding the structure [[xa]b]. Rule 1 will still be inapplicable in either theory since its structural description is never met. Rule 2 will be applicable under both theories since it would crucially involve material introduced on the current cycle. The final output under both theories would therefore be xrb.

Let us skip down to case four. Here, yet again, both theories will make the same predictions. x will pass through Stratum I with nothing eventful occuring. At Stratum II the affixes -a' and -b are affixed yielding the phonological form [[[x]a]b] on the relevant cycle. Now, since the application of Rule 1 would be permitted under both theories, we will apply it and derive the form xqb.

Finally, consider case three. Here is where the two theories differ. First of all, Consider theory A, which, as will be recalled, is the theory with no Bracketing Erasure. At Stratum I -a will be concatenated with x to yield the form [[x]a]. Passing on to Stratum II, we derive [[[x]a]b], which, as with the previous example will become xqb by the application of Rule 1, bleeding Rule 2.

Theory B, however, will treat this form differently. [[x]a] will be derived at Stratum I but at the exit of that Stratum the internal brackets

will be erased in accordance with the Bracketing Erasure Convention, yielding [xa] at entry to Stratum II. -b will then be affixed to yield [[xa]b]. Rule 1 will be inapplicable to this form, but Rule 2 will be able to apply yielding xrb.

Summarizing, the following are the predicted outputs of Hypothetical Language A for the two theories:

(82)

		Theory A	Theory B (LPM)
(i)	xab	xab	xab
(ii)	xa+b	xrb	xrb
(iii)	x+a+b	xqb	xrb
(iv)	x+a'+b	xqb	xqb

So the two theories are substantively different; they characterize non-identical intersecting languages.

The examples could be run through with equivalent results replacing phonological rules 1 and 2 with morphological rules, but the essential form of the argument would be the same. The question, then, is whether the predictions of Theory B (LPM) in this hypothetical case are consistently exemplified in natural languages or if there are cases where Theory A is a more adequate model. Case (iii) is obviously the crucial example; what we need to find in order to argue against LPM are examples of phonological or morphological rules which make reference to a morphological bracketing, and do in fact apply even when that bracketing is derived on a previous

stratum; LPM, with Bracketing Erasure, naturally rules such cases out in principle.

In fact, as Hargus (1985) argues, such cases do in fact exist in the Athapaskan language Sekani. I will not review all of the cases which she discusses but merely give the following example. In Sekani, there is a derivational process which is characterized by the combined prefixation of k'e and na, the latter of which is the customary/habitual prefix. Hargus terms this the "perambulative" derivation. An example is given below (from p. 313), where the numbers underneath the glosses refer to prefix positions:

(83)

k'è-na-n	-d	-zùt		===>	k'ènǫdzùt
per C	2sS	clf	V:skate		
2 5	12	13	stem		

(clf=classifier; 2sS=2 sing. subj.)

'you [sg] skate around'

Under certain conditions, this sequence of prefixes undergoes a rule which Hargus terms Perambulative Reduction, which she first formulates as follows:

(84) k'ènа ---> k'àn/ __ (C)[(clf) stem]

The result of this rule is seen in a form like k'àbeh 'he, she swims around' from k'è-na-d-beh (see p. 314 of Hargus).

Note that the environment is stated in terms of only partially

phonological information. In particular, Hargus argues it crucially refers to a morphological bracketing which contains an optional classifier followed by the stem. Now, Hargus argues elsewhere that the structure of the Sekani lexicon with respect to affixation is as follows:

(85)

Stratum I	aspectual suffix classifier prefix (position 13)
Stratum II	Positions 9-12
Stratum III	Positions 7-8
Stratum IV	Positions 1-6

According to this model then, and assuming Bracketing Erasure, the structure of the bracketing containing the classifier and the stem, which bracketing is derived at Stratum I, should certainly be invisible to the Perambulative Reduction rule, which, given that k'e (position 2) and na (position 5) are affixed at Stratum IV, must also apply at Stratum IV at the earliest. Nevertheless, Hargus argues that the formulation the rule given above is correct insofar as there is no way to make the rule work correctly without crucially referring to those brackets. The final version of the rule is given as follows:

(86) k'è-na ---> k'àn/ __ (C)[$_1$

We will return to the meaning of the subscript 1 momentarily.

For now, notice that this is precisely the kind of example we were

looking for: the real rule of Sekani behaves much as Rule 1 of Hypothetical Language A did with respect to morphological bracketings derived at previous strata. This, plus the other examples from Sekani which Hargus discusses, would seem to constitute strong evidence against the Bracketing Erasure Convention.

Nevertheless, Hargus has a different interpretation of the Sekani facts. She argues, following a suggestion by Kiparsky (1983), that Bracketing Erasure is exceptionable in the sense that particular morphological operations may specify that the brackets that they introduce be retained against the ravages of Bracketing Erasure. This is what the subscript 1 in the later formulation of the rule means; in Sekani, stems and classifiers introduce brackets which are exceptions to Bracketing Erasure, and they mark those brackets with some notational device like a subscript, so that Bracketing Erasure will know that it should leave them alone when the time comes for its application.

It will be recalled from Chapter 1 that Kiparsky's suggestion that Bracketing Erasure may be exceptionable was based on Bracketing Paradoxes. I argued there that this was not the right kind of motivation for assuming that Bracketing Erasure was violable since, in particular, such an approach would classify Bracketing Paradoxes as being not only "odd" from the point of view of morphological structure, but also from the point of view of phonology, for which I argued that there is no evidence. Hargus, on the

other hand, has provided precisely the right kind of evidence for demonstrating that Kiparsky's claim is in fact valid.

Nevertheless, one wonders, in the face of the Sekani cases, what the motivation for keeping Bracketing Erasure is altogether. Is the evidence for it so strong that we need to maintain it as a constraint on the lexical operation of phonological rules? I believe, in fact, that the evidence is not strong at all, and that the arguments for it are either faulty, that the facts can be explained by other means, or in fact are arguments for something else, namely the Strict Cycle. I turn to this question immediately.

One argument for Bracketing Erasure due to Mohanan (1982) has to do with the English rule of sonorant syllabification, which accounts for pairs such as hinder~hindrance. Mohanan observes that the rule only fails to apply with Stratum I affixes: note hindering, *hindring. This is captured if we assume that sonorant syllabification applies at Stratum II, and is stated as follows:

(87)

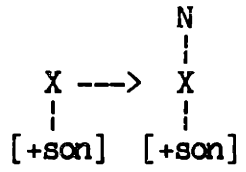
$$[+\text{ son}] \rightarrow \begin{matrix} V \\ | \end{matrix} \; / \; \begin{matrix} C \\ | \end{matrix} \_\_]$$

In a form like hindrance, the internal brackets will be erased around hinder before the rule has had a chance to apply at stratum II. Hindering, on the other hand will allow for this rule's application since, at Stratum

II, its structure will be [[hindr]ing].

Nevertheless, one could just as easily assume that the rule applies at Stratum II and is stated as follows (on the notation, see Levin (1985)):

(88)



This just says that an unsyllabified sonorant becomes a syllable nucleus. In the case of hindrance, the /r/ will already have been syllabified so (88) will not apply. It will apply to hinder yielding a syllabified /r/. Hindering is irrelevant since -ing will attach on the cycle after the rule has applied; given that (88) is a structure-building rule there is nothing within LPM to stop it applying to forms underived at Stratum II. (In the last section of the chapter I shall argue that this rule is a word-level rule and that -ing attaches to phonological words)

Hargus herself provides another supposed argument for the Bracketing Erasure convention, based on data from Mohanan (1982). She argues that in order for the stress and tone rules to apply correctly to subcompounds in Malayalam (see Section 4.2.2.3. above for a lengthy discussion of these facts), we must assume Bracketing Erasure. Recall that a subcompound such as [[mata][widweesam]] 'religion hatred' is stressed at the cocompound stratum as follows:

(89)
[[máta][widwèesam]]
L H H

Of course, the rule cannot be allowed to see the internal structure of the composition otherwise it might misapply to yield a cocompound stress on the form:

(90)
[[máta][widwéesam]]
L H L H

Bracketing Erasure guarantees this, of course. Nevertheless, we could also argue that stress rules apply to the maximal domain allowable, which would be the whole word in this case[29].

Another example is due to Kiparsky (1983a). These examples have to do with English zero-derived denominal verbs formed at Stratum II. Kiparsky notes that while such verbs may be formed from nouns which are themselves derived at Stratum I, such as to proposition, or to engineer, it does not seem to be possible with nouns derived at Stratum II:

(91) *to singer, *to freedom, *to promptness, *to alcoholism, *to sisterhood

The constraint proposed states that zero affixes cannot be added in English

---

29. This possiblity was pointed out to me by Donca Steriade.

to morphologically complex forms:

(92) *]X]Ø

This will derive the results mentioned above since a word like proposition will not be morphologically complex at Stratum II, where the constraint must apply, whereas a form like singer will be complex.

Nevertheless, I think that the generalization that this process is generally possible with Statum I nouns but never with Stratum II nouns is false. Consider the example to sticker (as in John stickered the wall yesterday.) This sounds like a fairly plausible verb to me, yet it looks as if it is derived from a morphologically complex form derived (by affixing -er to stick) at Stratum II. I am unable to think of further examples at the present time, so it may seem as if Kiparsky's point is valid. Nevertheless, notice that bona fide examples involving Stratum I suffixes do not exactly abound. Note the ungrammaticality of the following: *to grammaticality, *to religion, *to scientist, *to evasion.... If anything, it seems as if Kiparsky's constraint may in fact be applicable to affixed nouns no matter what their derivational history; forms like to proposition, to engineer and to sticker will simply be exceptional, in that case[30]. Of course, such a view would imply that

---

30. Marchand (1969: pp. 372-3) makes precisely this point. Furthermore he makes the following statement:

Bracketing Erasure is false.

Sekani provides another example, according to Hargus. There is a rule which Hargus calls Conjugation ə Deletion, and which she formulates as follows:

(93)

$$\text{ə} \dashrightarrow \emptyset \;/\; V \begin{Bmatrix} s \\ n \\ gh \end{Bmatrix} \begin{Bmatrix} [+\text{cnj}] \\ [+\text{mod}] \end{Bmatrix} \_\_\_\_ \; [(\text{clf})\ \text{stem}]$$

This rule applies, as Hargus puts it, "when no prefixes of the subject (position 12) or mode (position 11) prefix positions intervene between the conjugation or mode prefix and the optional classifier or verb stem." The following two examples illustrate the point, with the target vowel being underlined:

(94)

a.

na -sə -s -d -kwi ==> nasəskwi
rev-cnj-1sS-clf-V:vomit
5 10 12 13 stem
'I vomited'

> Another reason seems to be still more important. Many of the nominal suffixes derive substantives from verbs, and it would be contrary to reason to form such verbs as arrival, guidance, improvement, organization when arrive, guide, improve, organize exist.

There may well be something to this; note that Marchand's constraint is similar to morphological blocking.

b

na -sə -d -kwi	==> naskwi
rev-conj-clf-V:vomit	
5 10 13 stem	
'he, she vomited'	

In the first example the first singular subject prefix intervenes and blocks the rule's application, whereas in the second example no such prefix intervenes and the rule applies deleting the _ of the conjugation prefix.

Why do position 11 and 12 prefixes block the rule's application whereas the classifier prefix does not? Hargus suggests that since classifier prefixes are added at Stratum I (see above) and positions 10, 11 and 12 are all added at Stratum II, that application of Bracketing Erasure between the two strata can explain why the Stratum I prefix does not block the rule's application, whereas Stratum II prefixes do; the latter will be visible whereas the former will not. Nevertheless, one could presumably also just say that the rule fails to apply when phonological material intervenes between the schwa and the constituent [(clf) stem]. This is exactly Hargus' informal statement of the rule above, but it is not obvious that it is any less explanatory than the more sophisticated explanation making use of Bracketing Erasure.

The final example comes from Malayalam and is discussed in Mohanan (1982, pp. 52-3). Mohanan notes that a compound such as yakṣakinnarākkuuṭṭam 'a group of Yakshas and Kinnaras,' derived from yakṣan

'Yaksha,' kinnaṟan 'Kinnara,' and kuuṭṭam 'group,' exhibits gemination on the initial /k/ of the last member according to the regular rule of Gemination in subcompounds. The question is why the /k/ of kinnaṟan does not also geminate. Mohanan gives the following derivation:

(95)

Cocompound	[[yakṣan][kinnaṟan]]	compounding
	[[yakṣa][kinnaṟan]]	nasal deletion
	_____	vowel sandhi
Subcompound	[[yakṣakinnaṟan][kuuṭṭam]]	compounding
	[[yakṣakinnaṟa][kuuṭṭam]]	nasal deletion
	[[yakṣakinnaṟa][kkuuṭṭam]]	gemination

Gemination, which Mohanan gives as follows (p. 40):

(96) Onset ---> /\ % _____][

applies at the Cocompounding Stratum to the initial /k/ of kuuṭṭam, which is adjacent to a left bracket preceded by a right bracket (gemination is a mirror-image rule.) It does not apply to /k/ in kinnaṟan because the internal brackets of yakṣakinnaṟan have been deleted in accordance with the Opacity Principle (BEC). Nevertheless, it is also clear that the Strict Cycle Condition would rule out this application since nothing in yakṣakinnaṟan is derived on the cycle at which kuuṭṭam is added. Hence these data really just argue for the Strict Cycle condition, rather than for the BEC.

This concludes the examination of arguments for the Bracketing Erasure

Convention. I have shown, I believe, that the evidence for such a convention is weak and it is therefore questionable that the application of the phonology to lexical constructions is constrained in this fashion. Needless to say, if Bracketing Erasure is wrong, there is one less argument for distinguishing lexical phonology in particular from, say, its application above the word level.

We next turn our attention to cyclicity.

4.3.1.2 On Cyclicity in Lexical Phonology.

There are two questions which I would like to examine in this section. The less important one is whether lexical rules need be cyclic The more important question from my standpoint is whether post-lexical rules may be cyclic; Lexical Phonology has claimed that they cannot be (see section 4.1.1.) but it is far from obvious that this is in fact the case. The answers to both of these questions will be couched in terms of a short review of the available literature on the subject.

The first question, then, is: must phonological rules be cyclic? The answer to this question is certainly unequivocally no. It is not much of an issue that, as Kiparsky (1983b) discusses, a rule such as the /n/ deletion rule of English, which is responsible for the following alternations (from Kiparsky's exx. 8), is acyclic:

(97)

a. damn+ation damn+able, damn+at+ory, hymn+al, hymn+ody hymn+ology

b. damn#ing, damn#s, damn#, hymn#ing, hymn#index

A phonological rule such as the following (= (9)) would account for it:

(98) n ---> ∅/ [+ nasal] __]

However, notice that its application in a word like damn is already a violation of the Strict Cycle Condition since damn is an underived environment. A rule which does not obey the Strict Cycle cannot be cyclic; within current phonological theory it is taken to be the case that a rule is cyclic if and only if it obeys the Strict Cycle. Hence (98) is not a cyclic rule. That it must be lexical, at least under any reasonable understanding of that term, is indicated by the fact that it occurs before the affixes which, in Lexical Phonology, would be added at Stratum II. So lexical rules can be non-cyclic.

Of course, it is a reasonable question to ask as to whether one can restrict the domain of non-cyclic lexical rules; if any rule could be non-cyclic at any time, most or all of the content of Strict Cyclicity would be drained insofar as whenever one came across an example which seemed to require the added power of violating the Strict Cycle Condition, one could merely say that the phonological rule associated with the particular example was non-cyclic. But phonological theory is not contentless if we can restrict the domain of such rules to particular

levels of representation and if furthermore they are restricted in the way that they apply at that level then we may be saying something substantial. Indeed, Kiparsky (1983b) suggests that non-cyclic lexical rules must be restricted to the word level, which is Stratum II. This claim is also made by Halle and Mohanan, (1985). Also, see the discussion of stem/word phonology in section 4.2.2.

Let us turn now to the second main question to be addressed in this section, which was whether cyclic rule application is unique to the lexicon. The answer, apparently, is in the negative, and this is something which seems to have been known for quite some time.

Given that cyclicity is due to the fact that phonological rules intermingle with processes of composition--i.e., affixation more broadly construed--there is certainly no reason why we should expect that the claims of LPM should be correct, and there certainly exist counterexamples, which I shall now briefly mention.

The most famous set of cases comes from Catalan and is due to Mascaro (1976). Mascaró discusses the Catalan rule of Glide Formation, which he formulates as follows:

(99)
$$\begin{bmatrix} +\text{syll} \\ +\text{high} \\ -\text{str} \end{bmatrix} \text{----> [-syll] / [+syll]\#}_{0}\text{\_\_\_}$$

($\#_0$ stands for any number of boundaries.) So, an unstressed high vowel is turned into a glide after (any number of boundaries preceded by) a vowel. Mascaró argues that this rule applies cyclically and shows also that it applies across words as in pruduiráwksidəsyó 'it will produce oxidation,' which he derives as follows:

(100)

[[[pruduí]rá] [[[ɔ́ksid]á]syó]]

| | | |

lexicon, nevertheless it must apply non-cyclically in the post-lexical component. More recently, Dresher (1983) argued for the cyclic application of a post-lexical rule in Tiberian Hebrew. The reader is referred to his paper for the arguments.

In summary, we have argued for two points: First of all, as is fairly uncontroversial anyway, not all lexical applications of rules are cyclic; there must exist rules which apply non-cyclically though lexically at the word level. Secondly, we cannot expect that post-lexical rules should in principle be exempt from cyclicity. In effect then, cyclicity is not a special fact about lexical as opposed to non-lexical phonology. We turn now to the issue of Structure Preservation.

4.3.1.3 Structure Preservation.

Structure Preservation is a remarkably hard thing to say anything about insofar as the only formalizations of which I am aware, and which are due to Kiparsky, are stated in such a way that it would be hard in principle to falsify them; we can falsify particular interpretations of them since, as I shall suggest, at least some interpretations of them are, in fact, false. Still, the most general possible interpretation seems to say very little which is contentful; largely because of this, the following section will probably seem the most speculative so far.

Consider the following discussion from Kiparsky (1983a):

The second system governing lexical representations comprises conditions on what feature values may be marked. For example, in English voicing is distinctive for obstruents but not for sonorants. We express this by a marking condition which prohibits voicing from being marked on sonorants in the lexicon:

(16) $* \begin{bmatrix} \alpha \text{ voiced} \\ + \text{ son} \end{bmatrix}$

A language in which voicing is non-distinctive also for obstruents would have the marking condition

(17) *[α voiced]

By structure-preservation I mean that marking conditions such as (16), (17) must be applicable not only to underived lexical representations but also to derived lexical representations, including the output of word-level rules.[Footnote: We may wish to weaken this in various ways, e.g. by allowing marking conditions to "turn off" at some level of the lexicon, like phonological ruels, but I shall not investigate this possibility further here.

By assuming marking conditions I do not mean to claim that the learner assumes that all features are available for marking unless the language has specific evidence to the contrary. The reverse would be closer to the truth: the learner begins with the maximal restrictions and relaxes them only when he has to. Moreover, universal grammar will constrain this process by a hierarchy of features which defines their accessibility to marking.]

Structure-preservation greatly contributes to the restrictiveness of the theory since it determines point-blank that any rule which introduces specifications of lexically non-distinctive features must be postlexical. Thus the various rules for aspiration, glottalization, intonational features etc. in English could not be lexical. From the viewpoint of learnability this is an important constraint because it means that the leaner does not have to fix the domain of these rules by checking their ordering or other properties.

Nevertheless, despite the apparent benefits to be gained from Structure

Preservation, I believe that there are serious problems in its interpretation.

One question which arises is how the marking conditions are determined; that is, on what basis do we determine a priori for some language that a particular set of feature markings is non-distinctive? One particularly strong and therefore interesting interpretation might be that a lexical rule may never introduce a feature which is not distinctive in underlying representations[31]. So, if it is never necessary to assume that feature F is marked for any particular value so as to distinguish the underlying representations of two separate morphemes, then we could introduce the marking condition *[αF] which would guarantee that no rule in the lexical phonology would ever have access to that feature. Unfortunately this stronger hypothesis cannot be maintained, at least in the light of the results of Mohanan and Mohanan (1984), who argue that Malayalam distinguishes only five places of articulation for consonants underlyingly--i.e., bilabial, alveolar, palatoalveolar, retroflex and velar--and seven places at the output of the lexicon--i.e., bilabial, dental, alveolar, retroflex, palatoalveolar, palatal and velar. For nasals, in fact, only three underlying specifications seem necessary, namely bilabial, alveolar and retroflex, whereas the full seven are

---

31. Or, at least in underlying representation, plus whatever redundancy rules a la Archangeli (1984a) may be necessary for the language.

instantiated at the lexical level, the output of the lexicon. In the strongest sense, then, such a system is massively non-structure preserving.

Of course, there are tricks which one can use to maintain this stronger sense of Structure Preservation, and yet have a phonological system which seems at first sight to massively violate it. Such systems have been proposed; consider the following account of Breton consonant mutation as proposed by Carlyle (1983): In Breton, as in Welsh (Sproat, 1982, and see also the short discussion in chapter 1), consonants undergo various mutation processes, processes which would appear to be lexical rules by any criteria. The underlying inventory of consonants in Breton must at least distinguish the following segments: /p t k b d g f s z x m n l r/. Probably some of these, in particular the coronal fricatives, do not need to be as specified as I have implied here: for instance, we probably do not need to specify /s z/ as [+ strident]; in any event, this is beside the point for this discussion. What is to the point is that the phonological rules of mutation affect these segments in ways which are potentially counterexemplary from the point of view of the strong view of Structure Preservation. But first let us consider how this view of Structure Preservation might actually buy us something. The rule of spirantization takes the segments /p t k/ and converts them to the corresponding

fricatives, which turn out to be /f z x/[32]; /b d g m/ are unaffected by the rule. Now, Carlyle assumes (with Lieber (1983a); Sproat (1982), for Welsh) that the segments /p t k b d g m/ are underlyingly unspecified for the feature [continuant], in contrast with /f s z x/, which are marked [+continuant]. Furthermore, she takes the rule of spirantization to be a rule linking the specification [+continuant] to a consonantal matrix. In /f s z x/, of course, this will apply redundantly whereas in /p t k/ it will convert these to the corresponding continuant. Why does the linking not apply to /m b d g/[33]? Carlyle argues that the strong form of Structure Preservation would prevent this since there are no voiced continuants in the underlying inventory of Breton segments. So far, then, Structure Preservation seems to be buying us something in explanatory power since it constrains the application of a phonological rule in such a way as to allow the maximally simple statement of that rule; in particular, we would not have to state the fact that spirantization applies only to voiceless segments since this follows from Structure Preservation.

Turning now to lenition, we run into complications. Lenition converts

---

32. The voicing of /z/ is taken care of by a later rule.

33. As we see below, /m/ <u>does</u> lenite.

the segments /p t k b d g m/ into /b d g v z x v/[34]. Assuming, as Carlyle does, that the difference between voiced and unvoiced segments underlyingly is that the former are marked as [+voiced], and the latter are unspecified for voicing, there is no problem with the voicing of the voiceless consonants: we may simply assume that lenition links the feature [+voiced] to the segment. This will not violate the strong form of Structure Preservation. But how are we to turn a voiced consonant into a [+continuant] while still maintaining that: (i) there are no underlying voiced continuants; (ii) the strong form of Structure Preservation is correct; and (iii) consonant mutation is a lexical rule? In fact, Carlyle assumes that there is a feature [-tense] which is added to all lenited segments lexically, and that there is a rule converting [+voiced,-tense] to [+continuant] postlexically. We thus seem to have circumvented the condition on Structure Preservation but only by what amounts to "sneaking in the back door." It seems as if we genuinely can have both Structure Preservation and apparent violations of it in the same system[35]. Still, the criticism of such an approach must really be directed not at this analysis specifically but at the theory of phonology in general: Insofar as

---

34. Again, the devoicing of the velar fricative comes about by a later rule.

35. Note too that adding [-tense] to lexical representations could be taken as a violation of Structure Preservation given that [tense] is not underlyingly distinctive.

phonology is abstract and phonological features do not correspond in any simple way to articulatory features--a conception which has underlain most of the work in Generative Phonology and appears to have much to recommend it, it is not clear how we can rule out such analyses in principle: We could certainly imagine a marking condition for Breton which says *[-tense], but we would then be committed to having such conditions for any feature which one could come up with. Yet without such constraints Structure Preservation, even in its stronger sense, becomes virtually contentless.

In any event, given the Malayalam evidence, this stronger interpretation seems wrong anyway. What is the weaker interpretation? The weaker interpretation is that lexical phonological rules may not introduce features which are non-distinctive at the lexical level, i.e., at the output of the lexicon. This then, claims that no phonological rule may apply in a lexical derivation which would introduced a feature which would be non-distinctive of words. It is worth noting that even this weaker constraint may be wrong in its strongest sense: Hargus (1985) notes the following (p. 62):

> Rounding is predictable (i.e., non-distinctive) for [+high, +back] vowels in Sekani, suggesting that the feature [round] is not marked for high, back vowels. However, the output of Vocalization is a high, back, round vowel. Thus the prohibition against marking high, back vowels for any value of [round] cannot hold throughout the Sekani lexicon.

Thus, the suggestion which Kiparsky makes in a footnote (see above) to the

effect that a marking condition, like a phonological rule, may opt to "switch off" at some point in the lexical phonology, may well be more accurate. So even the weak version of Structure Preservation may be too strong.

Even given the weaker version, however, there still remains the question of how we determine the right set of marking conditions for a language, or, in other words, what, precisely the lexically distinctive features are. In fact, let us examine this question from the point of view of a language learner. If a child is learning English there will at some point be the determination that aspiration is non-distinctive. That is, there will never be a reason to contrast words on the basis of aspiration[36]. In fact, it is reasonable to assume that children learn this fact about English before they know much, if anything, about the lexical phonology of that language. So, at some point the learner will set up marking conditions such as *[αspread-glottis]. But notice that the rules introducing such features, which we may fairly well assume are learned as early as the fact that the features are non-distinctive, will have been learned <u>before</u> the setting up of such marking conditions. Needless to say, then, Structure Preservation does not aid at all in acquisition since such marking

---

36. It goes practically without saying that the learner could not know this about English beforehand since it could be exposed to Korean or Ancient Greek where the contrast is indeed relevant.

conditions must be set after the fact. This is not to say that there is no theory of phonological markedness (for which see, Kean (1974)), which will aid in the acquisition. But such markings are language specific and would appear to be set up after exposure to evidence for the relevant rules, and, crucially, before much if anything is known about the lexical phonology.

If this discussion is correct as I feel it is, it is interesting to consider the following point about the history of generative phonology. Chomsky and Halle (1968), following Chomsky (1964), and Halle (1959) among other works, eliminated the taxonomic phonemic level of representation which was current in structural linguistics; there seemed to them to be little motivation to keep it and there were certainly many theoretical problems associated with it. Nevertheless, this has always seemed somewhat wrong. More recently Lexical Phonology has reintroduced this level of representation as the lexical level. Mohanan claims, for instance, that the idea that the taxonomic phonemic level had something to offer was his major motivation for developing Lexical Phonology. He does in fact equate the level of taxonomic phonemics with the lexical level, but notes the following (p. 150):

> All the classical arguments against the taxonomic phonemic level of representation.....have been essentially arguments against a specific level of representation satisfying conditions like biuniqueness, local determinacy, invariance, and linearity. Since the level of lexical representations does not obey any of these conditions, the objections do not apply to the lexical level.

In particular "Taxonomic Phonemics focussed on the restrictions on the mapping between levels II and III [the phonemic and the phonetic: RS], while Lexical Phonology concentrates on the mapping between levels I and II [the underlying and the lexical: RS]." Of course, despite all this we still need to ask what the lexical level buys one. Mohanan claims that it is the level which is accessible to various psychological processes such as speech errors and language games; I shall be critiquing this view in the final chapter, though I should say now that I think that it is largely correct. But what does it buy the theory of phonology itself? What it buys, it would seem, is the determination of what rules may and may not be lexical and this in turn comes down to the weaker interpretation of Structure Preservation which we have been discussing. But while the Lexical Phonologist may well be concerned not with the relationship between the lexical and phonetic levels, it is patently clear that the language learner must first be primarily concerned with those levels. It is obvious that the learner, who starts off knowing nothing about the lexical phonology of the language to which it is exposed, cannot be concerned with the relationship between the lexical level and the underlying level since the latter is presumably unknown to it and can only be determined after a lot of acquisition--many years in the case of English latinate morphology.

So, while Lexical Phonology may well allow us to have the lexical level, and while this may buy us something in the determination of the phonological system and in psychological reality, we have nevertheless

reintroduced a level which apparently must be learned from the bottom up--more or less, in fact, as a Taxonomic Phonemicist might "acquire" the phonemics of an unknown language. And while this level may in fact be necessary, contra Chomsky and Halle, it seems that it cannot be viewed as a special benefit of Lexical Phonology since that theory does not apparently make the level any less problematic in the long run.

As promised, this section has been speculative. Nevertheless, it does seem clear at least that Structure Preservation, which seems like desirable constraint, is nevertheless not unproblematic in its interpretation.

In summary, I have argued in this section that Bracketing Erasure is of questionable value and motivation, that cyclicity is probably not a characteristic of lexical rules per se, and that Structure Preservation, and the lexical level which it implies, is not any more or less problematic in Lexical Phonology than in other imaginable systems which do not make use of the constructs which LPM uses. In the previous section, I argued that LPM, as a theory of word formation is of little content, at least in part because of the necessity of the Loop, and that doing away with the strata as "explanations" for apparent morphological ordering constraints not only forces us to look deeper into what is really going on in the morphology of languages, but is actually potentially successful in making strides in understanding this topic.

It seems then, as if not as much is explained as we might like from a

theory which claims that morphophonological, morphosyntactic and morphosemantic rules are all neatly bundled up in a highly organized component of the grammar termed the Lexicon. In particular, there seems to be no evidence that phonological rules behave "in the lexicon" in any ways massively different from the way in which they behave in the post-lexical component, beyond the cases where specific affixes make specific demands that particular rules apply or not apply. In the next and final section I shall very briefly sketch a theory of (morpho)phonological organization which is certainly not original but which would seem to afford an interesting view into the workings of phonology.

4.3.2 A Sketch of an Alternative to Lexical Phonology.

I turn now to the discussion of an approach which is an appealing alternative to Lexical Phonology, namely what I shall call the stem/word approach. As far as I can tell, there are two methods currently invoked to capture the fact that different phonological affixes have different classes of entities to which they can attach. One method is stratum ordering which is the method used by Lexical Phonology. The other method is to make a distinction between morphophonological entities such as stems and words and assume that various morphophonological processes select for one or the other of these. The latter approach is taken by researchers such as Selkirk (1982; 1984), Aronoff and Sridhar (1983), and Guerssel (1983, 1985); the reader is referred to these sources for expositions which will

be more complete in some respects from the one to be given here. In this section I shall discuss the advantages of such an approach over Lexical Phonology and I shall give a sketch of how it would tie in with my approach to morphosyntax.

4.3.2.1 Advantages of Stem/Word Morphophonology.

A stem/word theory of morphophonology claims that there is a limited inventory of morphophonological categories and that affixes select among those categories much as an affix may morphosyntactically select for a syntactic category such as verb or noun. In a language like English there would be two such types of entities, namely stems and words; for a language like Arabic it might also be useful to distinguish roots insofar as triliteral and quadriliteral roots play an important part in the morphology of that language (McCarthy, 1979). We may assume that this is a marked option, however. But what are stems and words? Stems I take to be the form in which basic lexical entries are listed. Furthermore, stems may be derived, as we shall see below, from basic stems plus affixes which select for stems. Words are the minimal units which have phonological independence—i.e., that can occur in phonological phrases or in citation form. Of course, these are purely functional definitions; I know of no independent criteria for determining what a word is as opposed to a stem cross-linguistically.

Basic lexical entries are listed as stems. Stem-level affixes selectively attach to stems. For English, such affixes would comprise the non-stress neutral affixes such as -ity, -tion, and so forth. Word-level rules may apply to stems at some point to convert them into words. Word-level affixes attach to phonological words. Affixes such as -ness, -dom, -hood, -s, -ing and so forth, are word-level affixes in English. Phonological rules, like phonological affixes, will be listed as applying either to stems or to words; the only exceptions to this will be rules which apply with specific affixes, and which must therefore apparently be listed with those affixes, more or less as part of the phonological "spell-out" of those affixes. A possible example of this in English would be a rule like Velar Softening which seems to be more or less a spell-out of the feature [+latinate].

How does this all work? Take a simple case like the by now very familiar word ungrammaticality. Skipping for the moment the question of how we derive the actual morphophonological bracketing from the set of options given by the Mapping Principle, a topic to which I turn in the next subsection, we will proceed as follows: grammatical will be listed as a stem[37]. Since such things as stress are predictable by rule in English, it

---

37. Needless to say, it may be arguable that even the form grammatical is morphologically complex. This is an unimportant point for the present discussion, however.

will be the case, among other things, that the stem is not listed with markings for metrical structure, unless, of course it is exceptional in that respect in which case it will have some metrical information listed. The entry for this stem would thus be something like the following: [gramatikal].

Now, as Kiparsky has convincingly argued (1983a), it is possible to view cyclic phonological rules not only as applying cyclically, but also applying so as to fill in features on lexical entries. In this capacity, structure building rules such as stress will apply to fill in features which are lexically unspecified; structure changing rules cannot apply because of the strict cycle (Kiparsky, 1983b). Application of stress to grammatical yields [gram átikal]. Now, -ity is listed as a stem-selecting affix (and is also marked as requiring a [+latinate] base.) As such. it can attach to grammátical and will do so.

There are a couple of points to make at this juncture. Say that when a stem-level affix attaches to a stem it always produces a stem. That is, there are no affixes which will attach to stems and thereby produce words, since the latter must be derived by further word-level phonology. If this is the case then the composite form grammaticality will also be a stem. Given this, we can enforce the reasonable requirement that, as a stem, it must look like a stem in that stem-level phonology must apply to it, including such phonology as fills in redundant features or structure, and

that, from the point of view of its segmental composition, it must also look like a stem in that it must in principle be enterable in the stem lexicon and hence must choose solely from the vocabulary of segments allowed by that level of representation. If this is correct, then we will have derived Cyclicity and the strong interpretation of Structure Preservation from the same source, namely that a stem produced by an affixation process must nevertheless look like any other stem in terms of its phonological structure.

This conception has at least one definite advantage: it is no longer necessary to build cyclicity into the model in quite the way that it is apparently stipulated in LPM: cyclicity will be derived from affixation coupled with a well-formedness requirement on representations. Notice too that we have linked Structure Preservation to cyclicity in a special way. We have claimed, possibly falsely and definitely falsifiably, that no cyclic rule which applies at some level, say the stem level, may change the inventory of segments which is available at that level. So if P denotes the set of segments specifiable in lexical entries of stems for some language L, then it will follow that all stems whether basic or derived must pick exclusively from P.

However, non-cyclic rules will not be constrained in this way. In particular, rules which convert stems into phonological words will not be required to limit themselves to producing outputs which select only from

the underlying set of phonemes. Rather there will be a second set of phonemes, possibly, though not necessarily, properly containing the first, from which well-formed words are formed. We will need, then, two alphabets, one which is given at the (stem) level of underlying representations and one which is given at the word-level. The latter will correspond to the taxonomic phonemic level or the lexical level of Lexical Phonology. This is not really a new suggestion insofar as the latter level is certainly necessary in LPM to derive Structure Preservation, whereas the former is also necessary to check well formedness of underlying representations. It is interesting to note in this context that the only phonological systems which have, to my knowledge, been discussed in the literature, which are massively non-Structure Preserving in the strong sense are Malayalam and Celtic; in neither of these cases is it necessary to assume that the new segment creating rules are cyclic. So, Mohanan and Mohanan (1984) argue that there is no reason to suppose that the rules creating new consonant segments in Malayalam are cyclic. And in Sproat (1982) I argued that not only was Welsh consonant mutation -which, like that in Breton, would appear to pose a counterexample to the strongest form of Structure Preservation--not necessarily cyclic, but that it could not be cyclic. Note too that we might expect that word-level rules are universally acyclic. The correctness of this assertion is suggested by Halle and Mohanan's (1985) discussion of English and Vedic, in both of which the word-level is acyclic.

The other point to be made about word-level rules which apply to phonological stems to create phonological words is that, insofar as they are not cyclic, they are unconstrained by Strict Cyclicity. This of course has been noticed before about word-level rules. Kiparsky, for example, defines Strict Cyclicity more or less in these terms. Note the following definition (Kiparsky, 1983b):

(101)
SSC: If W is derived from a lexical entry W', where W' is nondistinct from XPAQY and distinct XPBQY, then a rule A--->B/XP__QY does not apply to W until the word level.

In the case of grammaticality nothing of any great interest happens in the conversion from stem to word and un-, for instance, will be able to attach to the phonological word grammaticality. A more interesting case is a form such as damning. Here, the unsyllabified /n/, which is listed as part of the lexical entry of the stem and shows up in such stem-level morphological forms as damnation, is deleted by the word-level rule which we discussed in a previous section. This will be a prerequisite for the attachment of a word-level affix such as -ing. In this sense, following Aronoff and Sridhar (1983), we can say that the difference between word-level affixes and stem-level affixes is that the former produce a "closure" on their bases--i.e., they force them to cease to be stems and become words--whereas the latter fail to do so. So damning shows up as /dæmIŋ/ simply because of the fact that -ing must attach to a phonological word.

It is worth noting at this point that a stem/word approach to

morphophonology is more restrictive than LPM insofar as the latter predicts an arbitrary number of levels of morphophonology whereas the former would claim that in principle there could only be two; this point is also made in Aronoff and Sridhar's paper. Of course, LPM could be more restrictive if it were to constrain the number of levels universally.

Note also, again following Aronoff and Sridhar, that a stem/word approach does not predict ordering of phonological affixing. This is a desirable consequence since, as they argue (following Aronoff, 1976), there are facts from English morphology which are problematic for a stratum ordered theory such as LPM. Consider forms such as parsability and compartmentalization. In many respects -able and -ize behave like Stratum II affixes insofar as they both seem to attach to phonological words; note that they do not shift stress and clearly word-level rules such as Sonorant Syllabification and N-Deletion apply to them: hinderable versus *hindrable, damOable versus *damnable[38] and winterize versus *wintrize. Furthermore, what are for LPM Stratum II morphological processes, such as denominal verb formation, feed the affixation of -able: grandstandable. Yet apparently Stratum I affixes such as -ity and -ation attach to these forms. This fact could, of course, be handled in LPM by the introduction of a Loop between

---

38. Damnable with the /n/ pronounced is a word, but it has the very idiosyncratic meaning of 'despicable.' Hence it is not simply damn plus productive -able.

Strata I and II, but this would merely increase the adhocity of the model.

Within a stem/word approach these facts would be handleable as follows. Take the example with _-able_. _-Able_ would be listed as a word-level affix and it would be therefore required that it be phonologically attachable to a fully formed word. However, it itself would be listed as a (latinate) stem and stem-level affixes such as _-ity_ would attach to it and would hence form stems with the concomittant phonological changes such as stress shift. It is interesting to note, however, that the stress shift never affects the verbal base onto which _-able_ and _-ity_ attach:

(102)

compact	compactability
grandstand	grandstandability
butter	butterability

This suggests that the structure from the point of view of the _phonological_ interpretation of such forms is:

(103) (compact)(ability)

Indeed, Guerssel (1985) and Aronoff and Sridhar (1983) argue for precisely this phonological structure. We shall see in the next section that this structure will be deducible from the requirements of the affixes.

We turn now to the question of how this all fits in with the Mapping Principle and the theory of morphosyntax presented in chapters 2 and 3.

4.3.2.2 Stem/Word Morphology and the Mapping Principle.

In the first section of this chapter I argued that Lexical Phonology interpreted as a theory of the organization of the phonology and not as a theory of word-formation, is completely compatible with the assumption that the syntactic representation of words is different from the phonological representation. However, we have argued not only that LPM cannot be viewed as a theory of word-formation, but also that it is questionable how correct it is as a theory of the organization of the phononology, and we are now sketching an alternative to it. Therefore, it is incumbent upon us to show that there is as easy a fit between the latter approach and the Mapping Principle, as there is between LPM and the Mapping Principle.

This is not difficult to show. Consider an example like uncondemnability, which will have the following syntactic bracketing:

(104) [[UN [CONDEMN ABLE]] ITY]

The Mapping Principle will make the following general assertions:

(105) un^condemn^able^ity.

How, then, is the phonology to interpret this? That is, how is the phonology going to meet the requirements specified by the Mapping Principle, and still be able to phonologically interpret the string? The requirements will be as follows: -able must be adjacent to a phonological

word to its left, whereas un- must be adjacent to one to its right; furthermore -ity has to be adjacent to a (latinate) stem. Given these facts, we will be forced to proceed as follows: condemn will have to be converted from its underlying representation as a stem into a word. This will involve the stem level process of stress assignment, followed by the word-level process of Unsyllabified N-Deletion. This will yield an output like /kandɛ́m/. Un- is now free to attach, but can -able attach? Notice that if it does attach it will form the word uncondemnable, but since this is a word -ity could not attach. -Ity must therefore attach to -able before that affix is attached finally to uncondemn. In terms of cyclic domains, we have derived the following phonological structure for this word from the Mapping Principle plus the phonological requirements of the affixes:

(106) [[un [condemn]][abil ity]]

Thus we have derived the structure which has been argued for on similar grounds by Guerssel, and Aronoff and Sridhar.

In connection with the mapping from syntactic structure to phonological structure it is worth mentioning one other point here and that is the relationship, which has often been noted, between the phonological opacity of stem-level affixes, and their apparent semantic opacity on the one hand, and the phonological transparency of the word level affixes and their apparent semantic transparency. I have not encoded this in the theory

because I am not wholly convinced of the correctness of the generalizations. Certainly many highly semantically transparent operations such as inflectional morphology are often phonologically encoded in stem-level affixes: Latin presents fairly typical examples of this. Nevertheless, it does seem to be true that semantically opaque morphology is often encoded in stem-level phonological affixation. We might even want to state this as an implicational universal to the effect that if something is a semantically opaque process, then it will necessarily involve stem-level phonology.

This concludes the discussion of LPM which we began at the start of this chapter. I have argued that LPM is not a theory of word formation, and that therefore there really are no stratum-ordering effects in word-formation. Furthermore, although I showed that in principle LPM as a theory of phonological organization is compatible with the material discussed in the first three chapters, I argued that there are nevertheless serious questions with regard to how useful LPM is as a theory of phonology and how different lexical phonology is from post-lexical phonology. In this last section I discussed an alternative approach to morphophonology, one which I have also argued is compatible with the material in Chapters 1-3.

In the next and final chapter I discuss a few other issues that are raised by the approach to morphology taken in this thesis. Before we leave this discussion, however, it is worth noting that there is a residue of LPM which is still, I think, more or less correct. There does seem to be a sense in which phonological rules can "enter" the phonology with some affixation process, and "leave" it with another[39]. This, of course, is part of the content of strata in LPM. The effect shows up quite nicely in Sekani as discussed in great detail by Hargus (1985), where there seem to be very definite "breaks" in the sequence of affixes which delimit the domains of various phonological rules. I have nothing to say about this interesting phenomenon except that, while it is certainly consistent with LPM, the mere existence of domains of phonological rule application do not, in my opinion, warrant building an entire theory which is centered around this one point.

---

39. Thanks to Donca Steriade for discussion on this point.

Chapter 5

Some Final Points.

In this final chapter I wish to look at a few residual issues which are brought up by the foregoing discussion. First of all, I shall look at some further theoretical points: the nature of morphological blocking, the notion "component of the grammar," and the importance of the Projection Principle. The second section will be devoted to what may be termed "psychological issues:" Kohanan (1982) in particular has claimed that the lexical level--i.e., the level of the output of the lexicon in LPM--is an important level from a psycholinguistic point of view insofar as it is the level at which such things as language games, speech errors, and speech processing take place. I shall be examining the correctness of those claims. Finally, in the third section, I outline a future research program for morphology, under the assumption that morphology is not "done in the lexicon."

5.1 Some Theoretical Issues.

5.1.1 Morphological Blocking.

Imagine for a moment that the lexicon is simply an array of a (possibly infinite) number of pigeonholes. Each of these pigeonholes is associated with a particular concept or grammatical function. Imagine further that each pigeonhole may be filled with exactly one word which is appropriate for the particular concept or grammatical function associated with that pigeonhole. In this context morphological blocking as a principle specifically of morphology would make a great deal of sense; we would be making the rather natural claim that for any given slot which we need to fill, there would be space for exactly one item. More specific rules, insofar as they would be ordered before more general rules would always have a chance at filling the slots first and hence would always block the more general rules.

In the last chapter I argued that it was not generally true that affixes which are, in LPM, added at a later stratum than other affixes, are necessarily unable to block affixes which are added at an earlier stratum. In particular *un-* and *in-* seem to block each other in the sense that there are words which presumably have all the right properties to take, say, *in-*,

but would nevertheless appear to merely be marked to take un-. Putting that issue aside, however, it seems as if the pigeonhole concept outlined above, though appealing, is not quite right: One could certainly imagine, following Kiparsky (1983a) (among others), that every time a process of affixation is performed, the product of that affixation is entered as a lexical item; if by "enter as a lexical item" we mean to actually enter the form into some array of pigeonholes such as that just described, then the blocking effect falls out naturally. The reason this approach is not quite right, however, is the following: it seems as if blocking, or something which looks so much like blocking that it might as well be called blocking, occurs not only between items which arguably could be (or should be) stored, but also between items where only one of the pair--the construction which blocks--is arguably entered in the lexicon.

The best example of this point of which I am aware comes from Irish and has to do with the behavior of subject agreement in that language. It is a fairly well-known fact, and one which has most recently been discussed by McCloskey and Hale (1984), that Irish absolutely prohibits verbal agreement whenever there is an overt subject; this much Irish has in common with other Celtic languages such as Breton. What is unique to Irish (as also discussed in McCloskey and Hale) is the fact that a construction of the

form $V_{[-Inflection]}$ **Pronoun**[1], is used if and only if there is no inflected form of the verb corresponding to this particular syntactic construction. So, consider the following paradigm for the Conditional of the verb cuir 'put,' in the Ulster dialect (examples from McCloskey and Hale (p. 489)):

(1)

	Singular	Plural
1	chuirfinn	chuirfimis
2	chuirfeá	chuirfeadh sibh
3M	chuirfeadh sé	
F	chuirfeadh sí	chuirfeadh siad

Notice that some of the forms, namely the first and second persons singular, and the first person plural are fully inflected, whereas the other entries are composed of chuirfeadh, which is the unmarked form in this mood, followed by the appropriate pronoun--either sé 'he,' sí, 'she.' sibh 'you (pl.),' or siad 'they.' Now, it is ungrammatical in Irish to put a pronoun after an inflected form: *chuirfinn mé, where mé is the first person singular pronoun; pronouns or full NPs may only occur after the unmarked, uninflected form of the verb. This is explained by McCloskey and Hale's analysis and the reader is referred to their paper. However, what is also interesting is that when an inflected form exists, although the corresponding form with the uninflected verb followed by a pronoun ought,

---

1. Where Pronoun is the subject; Irish is a VSO language.

in principle, to be allowable, it is in fact always ungrammatical. Thus, *chuirfeadh mé 'I would put,' is completely ungrammatical though it is as syntactically well-formed according to McCloskey and Hale's principles as chuirfeadh sé 'he would put.' How is this blocking to be accounted for? Hale and McCloskey suggest the following:

> There are difficult questions of principle here, but rather than answer those questions, we would like to suggest that they contain a clue as to how the ungrammatical *chuirfeadh mé. . . is to be dealt with. We think that the necessary mechanism is that of morphological blocking--the phenomenon whereby the application of a morphological rule of limited productivity (e.g. the rule that derives went from go), blocks application of a more productive rule to the same stem (*goed).

I completely agree with McCloskey and Hale's assessment, but notice that there is a problem here. Certainly the blocking here is consistent with standard examples of morphological blocking: in general, blocking occurs between more specific processes and more general processes. But what cannot be upheld in these examples is the view that blocking consists of two morphological processes vying for a single location in an array, with the more specific process always winning out. The reason for this is that presumably a sequence of a verb followed by a pronominal subject is not a word, except perhaps phonologically if the pronoun happens to be a clitic. There is therefore no reason, at least under most conceptions of the

lexicon, to assume that such a composite is entered anywhere[2]. In fact, not only does blocking have nothing to do with the listing of both forms, the blocker and the blockee, but it also does not appear to be specifically morphological, since apparently a morphological form may block a syntactic composition.

Interestingly, Kiparsky (1983c) proposes a constraint which is very close to being right, generalizing somewhat. What he suggests is an "Avoid Synonymy" principle, stated as follows (p. 16):

(2) The output of a lexical rule may not be synonymous with an existing lexical item.

So far this is a statement about lexical rules, but the statement may be generalized:

(3) No (morpho-)syntactic composition may be synonymous with an existing lexical item.

This captures the entire class of morphological blocking cases, as far as I can tell: that is, as long as we have to assume, as I think we must, that specific forms are merely listed, then (3) will apply. It also captures

---

2. This is not to say that Verb Subject sequences, along with even larger syntactic chunks, are not stored physically in the brains of speakers of Irish; this is a different question however (see section 2 of this chapter for a discussion of this.)

the Irish facts: it is an idiosyncratic fact about the particular mood of the Irish verb paradigm that certain parts of the paradigm have specially inflected forms associated with them, whereas others do not: a form like chuirfinn[3] must be specially noted, therefore, as existing since this is not predictable on independent principles. So this special form will block, as McCloskey and Hale rightly suggest, the more generally available construction.

(3) has another effect, also suggested by Kiparsky for his version of the principle[4], and that is that it will block synonymy between two morphologically simple forms. The basic claim being made is that two constructions, whatever their complexity, cannot be synonymous. If there is a more specific construction--i.e., one which must be listed anyway--then that will always win out over a more general construction, which would normally mean the same thing, either by blocking it entirely, or by forcing it to take on a meaning which might well be within its semantic domain, but to which it would not normally be restricted[5]. What

---

3. Or at least X-finn; i.e., that this particular combination of person-number, mood and tense features exists as a verbal affix.

4. Both in the paper cited, and in Class Lectures, MIT, Fall 1982.

5. An example of the latter case (from Kiparsky, 1983c) would be drill which blocks driller from referring to the instrument and forces it to retain only the referent of 'person who drills.'

if there are two morphologically simple words? The "Avoid Synonymy" principle would, as Kiparsky argues, block such from being synonymous, since, given that one of the two must be entered with a particular meaning in order for there even to be a question of synonymy, the other will subsequently blocked from having that meaning. Given this, it is unsurprising that there are, at least in English, few convincing cases of synonymy among basic lexical items. I should note that "synonymous" here means more than just having the same denotation. Two words may well denote the same object, such as cop and policeman, but fail to be synonymous in that they have different uses, or different statuses in the language. So cop is more colloquial than policeman and hence has a different social status; the two words mean different things insofar as they impart different information about the social context in which they are used.

Still, the principle is so strong that people will often assume even before they know the exact meaning of two words, that they must actually denote different things. A fairly standard case would be something like ash and elm, which almost any adult speaker of English would know to be names for trees. Few people, however, would be able to reliably identify the referent of either of these terms, but there is nevertheless a very strong principle which would appear to be blocking the assumption that the two words mean the same thing. A principle such as (3) would accomplish this insofar as a speaker would know that whatever semantic value would end up being associated with one of the pair, the same semantic value would

simply be forbidden from being associated with the other member of the pair.

Blocking, we have argued, is not a specifically morphological principle; that is it does not constrain the application of specifically morphological rules, but rather applies much more broadly. This is, of course, entirely consistent with the approach to morphology which I have taken in this dissertation.

5.1.2 What is a Component?

In this thesis I have argued that the lexicon should not be considered to be a component of the grammar; that is, it should not be considered to be a separate "place" where one does word-formation. But what is a component anyway? I believe that the right characterization of the notion component is that it is a coherent subsystem of rules that forms a level of representation or relates that level to another level[6]. In the LF component, for instance, Binding Theory applies to check well-formedness of Binding Relations, and the Generalized ECP also applies to check the well-formedness of constructions involving empty categories. At D-structure, it can be argued that the theta requirements of lexical items

---

6. Thanks to Noam Chomsky for this particular way of formulating the notion.

in a construction are checked and that X-bar theory applies. And in the PF component phonological principles apply to interpret syntactic constructions; other principles such as Case-theory have also been argued to apply here.

But in what sense is the lexicon, qua word-formation device a component? The idea that it is such, and that there are specifically lexical principles which apply to constructions below the word-level--which constructions are internally invisible, at least to syntax--is a common one, I believe, and is by no means incoherent. One could well imagine a model of the grammar where the lexicon is a component orthogonal to all of the others, and which looks at words which are given (as atomic units) at the various other components of the grammar, checking via its own principles whether such words are well-formed. A schematic diagram of such a conception is given in Figure 1.

[INSERT FIGURE 1 HERE]

In this diagram the word which the lexicon sees as [x[y z]] is projected to the other components of the grammar as atomic units, represented here by the solid lozenges. So the lexicon "supervises," as it were, part of the grammar in that it is responsible for the well-formedness of words. The result is a fairly appealing conception of the interaction of independent components.

I have argued, however, for a different conception. I have pushed

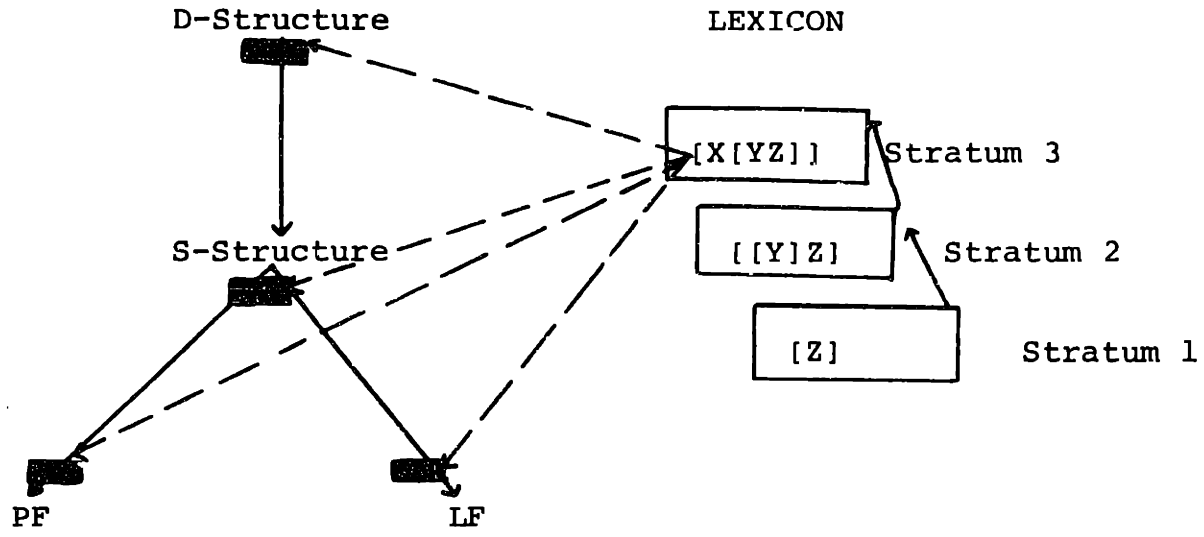

D-Structure
LEXICON
[X[YZ]]
Stratum 3
S-Structure
[[Y]Z]
Stratum 2
[Z]
Stratum 1
PF
LF

FIGURE ONE.

forward a model where morphology is distributed over the various components of the grammar, or at least in the syntax and in the PF component. That is, principles such as X-bar theory the syntactic Projection Principle, Case theory, Binding theory, Theta theory, and so forth apply to determine the syntactic well-formedness of words in general; such an approach was also taken by Fabb (1984), as has already been noted, but he limited himself to what he termed "syntactic affixation." On the other side of the fence, the phonology acts as interpretive component, interpreting the strings derived via the Mapping Principle from the syntactic representation of words; what kinds of phonological principles might be at work were discussed in the previous chapter. Under this conception, there is no word-formation component. There is however, a lexicon, and this will consist of what has come to be termed the "permanent lexicon," or the data structure containing all idiosyncratic forms in the language. This data structure interacts with the other components of the grammar by projecting the requirements of the particular lexical entries to the relevant components of the grammar. A diagram of this model of the grammar is given in Figure 2.

[INSERT FIGURE 2 HERE.]

In this diagram X' and Y' represent different morphemes which are listed in the permanent lexicon with their syntactic and phonological halves $\underline{X}$, $\underline{Y}$ and $\underline{x}$, $\underline{y}$, respectively. Each morpheme projects its requirement separately to

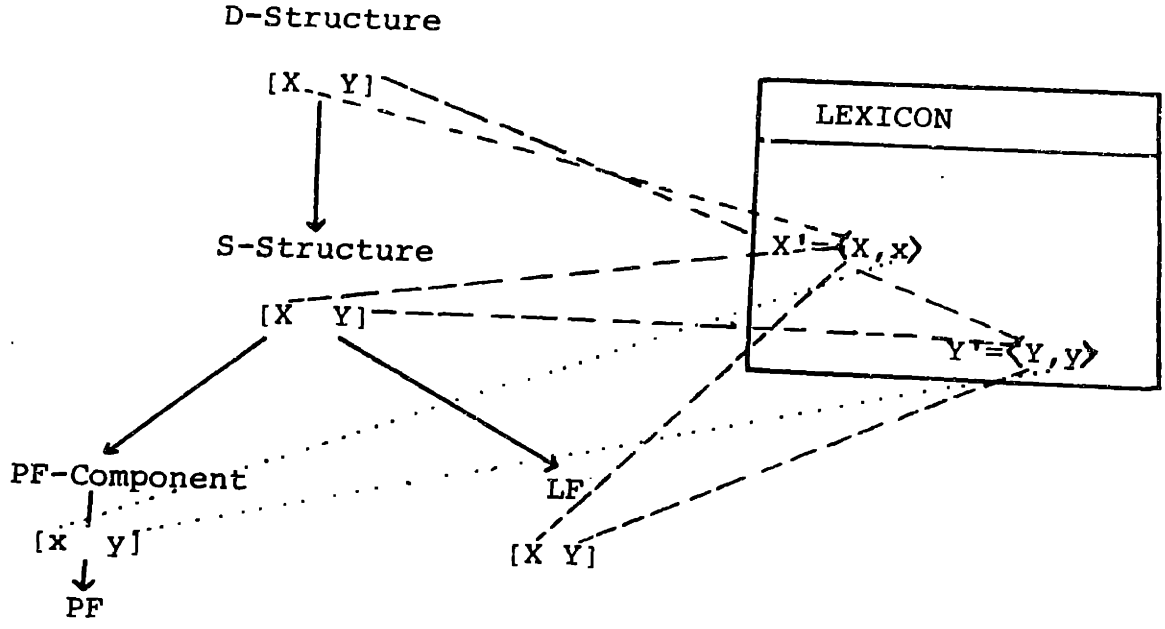

D-Structure
[X Y]
S-Structure
[X Y]
PF-Component
[x y]
PF
LF
[X Y]
LEXICON
X'=⟨X, x⟩
Y'=⟨Y, y⟩

FIGURE TWO

the various other levels of grammar. In particular, in the syntax the collocation of X and Y is checked for well-formedness and in the phonology the sequence xy will be interpreted. The lexicon itself is some sort of data structure which holds information about morphemes and idiosyncratic complex forms.

There are two things of note here. Firstly, if my approach is the correct way to view morphology, then we have come almost full-circle to the position of SPE with respect to what the lexicon itself is. Despite much research on the lexicon as word-formation component, however, I do not know of any convincing evidence that the SPE conception was wrong. In fact, the vast majority of the research on the lexicon qua word-formation component has focussed on a different question, namely how the phonology is organized, which is of course, a very interesting problem; but an answer to that question need not show that words themselves must be viewed as having all of their properties derived in one separate location of the grammar. A jump from the clearly correct idea that phonological affixation and phonological rules must be intermingled in a way first suggested by Pesetsky (1979) to the view that all facets of word-formation must take place "in the lexicon" involves, as far as I can see, completely fallacious reasoning. Needless to say, such reasoning has been taken seriously: Lieber (1980) argued, on purely morphophonological grounds that inflection must be done in the lexicon. Even assuming that her arguments are correct,

however[7], the fact that inflectional morphology behaves exactly phonologically like derivational morphology, tells us precisely nothing about the way that inflectional morphology should be treated in the syntax[8]. Nevertheless, Lieber's arguments have been the basis for many a claim that a theory such as LFG where the bulk of syntax is done "in the lexicon" must be more correct than other theories where this is not the case[9]. When the syntactic/semantic aspects of word-formation are separated from the phonological aspects of the same, as I think they must be, the arguments for a separate lexical component look much less convincing.

The second point is that what I have called the (permanent) lexicon, is by no means an uninteresting object. The reason that I did not refer to it as a list, but rather with the more vague term data structure is that the former implies that there is nothing particularly interesting about the relationships between the entries in the list, whereas the latter has no such implication. Indeed, as Rappaport and Levin (1985) and much other work in lexical semantics has suggested, there may indeed be a very rich structure to this database. Needless to say, this has not been the topic

---

7. See Thomas-Flinders (1983) for evidence that they are not correct.

8. There may be genuine syntactic arguments for not distinguishing these two types of morphology; see Marantz (1984b) on this question. However this is an entirely different issue.

9. See Bresnan (1982: p. 307), for instance, among others.

of investigation here, but I feel that much interesting work has yet to be done on the lexicon, viewed as a database of basic and idiosyncratic lexical items.

In describing Figure 2 above, I noted that well-formedness of words (and indeed phrases and clauses) could be determined by examining the requirements of the various component lexical items which have been projected from the lexicon. This notion of projection is a familiar one for syntax, but it is perhaps not so familiar for phonology. I turn to this point in the next section.

5.1.3 On the Projection Principle.

Let us repeat here the definition of the Projection Principle as given in Chomsky (1981):

(4) Projection Principle.

> Representations at each syntactic level (i.e., LF, and D- and S-structure) are projected from the lexicon, in that they observe the subcategorization properties of lexical items.

The Projection Principle, then, is taken to be a property of syntactic representations. In fact, I have argued (Chapter 2) that the Projection Principle not only holds of lexical items which are inserted into phrases, but in fact in general for any X^0 whenever it is the head of a construction and thus dominated by a node of category X:

(5)

X^n
|
|
X^0

(where n>=0)

This will include not only the phrasal instances where X is the head of a maximal projection, but also instances, which have usually been assigned to the lexical component, where X is not dominated by a maximal projection: such are the synthetic compounds discussed at length in Chapter 2. The Projection Principle will also be at work in morphology in assuring that at all relevant levels of representation, an affix which subcategorizes for an item of category X, will be sister to such an item.

What is generally held, however, is that the Projection Principle does not hold of PF. That is to say, if a verb such as see requires an object, then there must be some NP, if only an empty category, which is governed by see at D-structure, S-structure and LF. On the other hand, there is no reason to suppose that the Projection Principle requires that there be a representation for the object of see at PF. In fact, why should it? Since the Projection Principle is stated as holding over the syntactic requirements of lexical items, there is no reason why PF, which we take to be the output of the Phonological Component, should even be concerned with the syntactically relevant Projection Principle.

Nevertheless, I think that a form of the Projection Principle does hold in the PF-component, and that the Projection Principle as stated above should be generalized so as to account for this. Remember from Chapter I that we have envisioned a lexical entry as having the following form:

(6)

$$L' = \langle L, l \rangle$$

L', a lexical entry, is a vector whose components consist of L, the syntactic representation of the morpheme, and l, the phonological representation of the morpheme. Now, the standard Projection Principle will apply to L at any level in which that half is relevant--basically, D-structure, S-structure, and, we may assume, LF. But a projection requirement of a sort holds of l in that, in the PF-component, l is constrained to behave in a particular way. For example, suppose for a minute that l is listed as a phonologically bound morpheme--a phonological affix. In that case, certain phonological requirements will hold of l--due to the lexical properties of l--in that it is constrained to attach to certain kinds of phonological constituents such as stems or words, or at a certain stratum if we take a stratum-ordered approach to morphology. Even supposing that l is not an affix, hence phonologically unbound it must project properties of itself to the PF level of representation. In particular its representation at PF (which we may take, following Chomsky (1981), to be the level of phonetic representation) must be a composition of its lexical representation (i.e., what is actually listed as the second

component of L') with the phonological rules of the language, taken as a function from underlying to surface forms.

This point has been expressed, in fact, by Chomsky (1984 138-143), under the guise of the Full Interpretation Principle:

> We might express many of these ideas by saying that there is a principle of Full Interpretation (FI) that requires that every element of PF and LF, taken to be the interface of syntax (in the broad sense) with systems of language use, must receive an appropriate interpretation -must be licensed in the sense indicated. None can be simply disregarded. At the level of PF, each phonetic element must be licensed by some physical interpretation. The word book, for example, has the phonetic representation [buk]. It could not be represented as [fburk].

Chomsky goes on to note that there cannot be sentences of the form '[who] John saw Bill' where there is vacuous quantification, which is simply ignored; this is not a constraint on formal languages in general since statements such as '(Ax)(2+2=4)' have a perfectly straightforward interpretation by simply ignoring the quantifier. On the other hand FI seems to be a property of natural language.

Nevertheless, while it is not clear how the LF half of FI fits into this, it seems as if the PF half, extending the notion from "physical interpretation" to phonological interpretation, is really just a case of Projection applying in the PF component; fburk is not a possible representation for the phonological/phonetic half of the lexical entry BOOK' trivially because the Projection Principle would require that the extraneous segments be projected into the phonology. I therefore

generalize the Projection Principle as follows:

(7) Projection Principle.

> Representations at each level of Grammar are projected from the lexicon: representations at each syntactic level (i.e., LF, and D- and S-structure) are projected in that they observe the subcategorization properties of the syntactic component of lexical items; representations in the PF component are projected in that they observe the phonological requirements specified by the phonological component of lexical items.

We now turn from these theoretical points to some questions about the psychological status of the lexicon and of word formation.

5.2 Psychological Issues.

5.2.1 On Productivity, Listing and the Separation of Word-Formation from the Rest of Linguistic Knowledge: What is the Relationship between the Theory of Morphology and My Knowledge of Words?

It was Aronoff (1976), following in the footsteps of Halle (1973), who argued that the study of morphology is the study not of the actual words of somebody or other's lexicon, but rather of the possible words of a language. There is a good reason for this: you or I, as speakers of English know a great deal about the structure of words, but which words you happen to know and which ones I happen to know are highly idiosyncratic,

based on experience and therefore not of great theoretical interest from the point of view of linguistics: it is irrelevant to the study of morphology that I might know the word basidiocarp and you might not; or that I can never seem to use the word fortuitous correctly whereas other people do not have this problem. What is important is what a speaker of a language knows about the (productive) morphology of that language. That is, whether or not a speaker of English, for instance, would consider that a word such as oligarchical is a believable word of English even if he or she is not familiar with it.

Of course, there are many factors which contribute to such a decision, not all of them strictly within the domain of morphology. A classic case of the latter would be the non-existent but possible (monomorphemic) English word blick, which is certainly non-existent but certainly possible, not for morphological reasons, but for reasons having to do with English syllable structure.

Of course, despite the ideal models that a morphologist may construct, the fact will still remain that not only are the actual words of someone's lexicon idiosyncratic and hence uninteresting, but that what each individual speaker will know about the productivity of various affixes will also be highly idiosyncratic. For instance, for me, almost any Latinate sounding word which I either know to be or can believe to be an adjective can take -ity: basidiocarpicity, and blopiticity are both reasonable

sounding words to me; basidiocarpic is an adjective which I know, although I have never actually seen the -ity noun derived from it; blopitic does not exist for me, but if it did, I could well believe that it would be an adjective and that it would take -ity. Now, a speaker of English who is fortunate enough not to have spent the last two and a half years thinking about this kind of stuff is very likely not to concur; for him -ity affixation may not be productive at all, and may have to be listed with the few words which he knows which contain the morpheme. So, when a speaker of English asks, as I have been asked many times, "is that a word?", what would appear to be going on is that the speaker is simply unsure as to the productivity of the particular morphemes involved.

Of course, the comparable question with phrasal syntactic units would be absurd: no English speaker, for instance, would question whether snarfed the rutabaga cheesecake is a possible verb phrase of English. So, one might imagine nonproductivity to be the hallmark of the lexicon: non-productive processes might be said to take place within a separate component of the grammar from the syntax, which deals only in productive processes.

Well, in one sense, I agree: Unproductive formations must be listed under anybody's theory. That is, if any part of a linguistic composition is irregular, either phonologically, or because of syntactic or semantic irregularity, then this fact must be notated somewhere, presumably by

listing the irregularity. That place, I have suggested, is the permanent lexicon, for which I have reintroduced the term "lexicon" in its original SPE sense. On the other hand, it does not follow from this by any means that there is a separate word-formation component which deals among other things, in idiosyncracies. We do not need this notion. For one thing, there is much that would be considered lexical that is unquestionably regular in salient respects. Compounding, for instance, appears in theories of morphology which have been developed recently, but there is no question of irregularity here: Nobody doubts that nonce-compounds such as bog-dog, hog-log, rutabaga-demolition and scurvy-avoider are possible words of English. With respect to the second two, since they are synthetic compounds, there would be little doubt as to their semantic interpretation. With respect to the first, insofar as root compounds are often idiosyncratically interpreted (see Downing (1977)), there will not necessarily be agreement as to what they mean although syntactically no one will be in doubt. Of course I lied before when I said that there was nothing irregular about these formations since idiosyncratic interpretation would certainly fall under that rubric. But notice that the irregularity, as Downing convincingly argued, is not conditioned by linguistic factors anyway so it would not help one to have a lexicon qua word formation device so as to take care of this set of examples which in terms of syntactic structure are quite regular. So-called Stratum II affixes such as -ness enjoy a status similar to compounds in English: with shagginess, nurdiness,

weightiness, there is little question as to interpretation or well-formedness.

The fact of the matter is that productivity does not slice the pie correctly with respect to the division of components as many theories of word-formation would have it. Of course, one could back down from this position and claim that only idiosyncratic processes belong in the lexicon qua word-formation component. This is the position taken in Fabb (1984), where it is claimed, for instance, that affixes such as -ness and constructions such as synthetic compounds are not lexical but syntactic. Less productive formations, such as nouns derived from adjectives via -ity, would be formed in the lexicon. Still, I see no reason for making even this distinction. Let us grant, for the sake of argument, that -ity affixation is not productive. A word like grammaticality will therefore have to be listed in the permanent lexicon; this is a reasonable assumption insofar as either (i) an affix is productive in which case nothing formed from it need be listed, unless a particular form is idiosyncratic in some way, or (ii) an affix is not productive in which case all of its formations must be listed. Still, if we assume that anything at all is known about -ity as an affix--i.e. its phonological and morpho-syntactic properties[10],

---

10. And if we cannot at least assume this, then we are claiming essentially that a word like grammaticality is completely idiosyncratic and hence all of its levels of representation must be listed; it is not "formed" under anybody's story.

then there is right away no reason for assuming that it is formed anywhere; its properties can quite easily be checked in the syntax and phonology--to which components it must be projected by our assumptions, anyway--and word formation "in the lexicon" can be dispensed with entirely.

One other thing needs to be mentioned before leaving this topic. Aronoff has claimed as a fact about words the following (p. 18):

> ..words are peculiar, not only in that not all of those that should exist actually do, but also in that those which do exist do not always mean what they are supposed to mean, or even look like what they are supposed to look like. Words, once formed, persist and change; they take on idiosyncrasies, with the result that they are soon no longer generable by a simple algorithm of any generality. The word gravitates toward the sign.

There is, of course no doubt whatsoever that this is true of words, but is it true exclusively of words? I think not. Noun phrases, for instance, are also able to take on idiosyncratic meanings: lion's share certainly can be interpreted literally, but there is a preferrable interpretation for English speakers who know this idiom which is at best metaphorical and certainly not predictable. The prepositional phrase in the red once had a very transparent reading in the days when accountants would enter debits in red ink; now it is used to mean 'in debt' by people who have no knowledge of this tradition[11]. In languages such as French, Spanish, Italian, Welsh

---

11. Notice that it is really a frozen expression: *in some red, *in a red, *on the red, so it is not just an idiosyncratic reading of red.

and Arabic which do not have a productive compounding strategy, many things which are in English formed by compounding are full-fledged, but idiosyncratically interpreted noun phrases. French Salle à manger 'dining room' (literally 'room for eating') is a good example of this as is ystafell wely 'bedroom' (literally 'room of bed') from Welsh.

This is a hardly surprising fact really, and one which must be deeply rooted in psychology. Not only do speakers of languages physically store many words which have idiosyncratic interpretations, but they must surely also store much larger chunks, even ones which are perfectly transparent in their construction and interpretation. I doubt, for instance, that every time I hear someone say 'close the window' I need to compute from scratch the semantic structure of this sentence. I have heard it so many times that it almost goes without saying that I have it represented somewhere or other in my mind as a chunk, even though a perfectly analytic chunk. Evidence for this is that I can remember, for instance, that someone might have said this particular sentence to me on a certain day at a certain time and that there was a particularly weird way in which this person said it; perhaps it was sung to the tune of the Little Fugue in G-minor. Given that I can remember how this sounded (and what it meant, given that I also remember going to the window and shutting it), it seems unlikely that I would, every time I recall this event, basegenerate the sentence, perform whatever syntactic and phonological operations are required and then run it through a conveniently available Little Fugue in G-minor filter so as to

reconstruct the recollection. More plausibly, I remember the whole thing as a chunk. So, I think that there is reason to believe that larger units than words must in fact be actually, physically entered, independently of their productivity. From the point of view of psychology, however, this puts words along with larger chunks on the same basis with respect to one point, namely storage. But if larger units can be stored, then they can also, like words, be stored with particular idiosyncratic meanings associated with them. Of course such meanings will not generally block the productive compositional interpretation: it is possible, though difficult except in contexts which force the literal interpretation, to interpret kick the bucket or chew the fat literally. But this is also true of many lexical "idioms." Transmission has a very special meaning associated with it: it refers to a part of an automobile. But this does not in general block the other, productive, nominalization, as in 'The transmission of information via satellite.' I should note in closing, that the point about syntactic chunks being stored is not original: it has been made by Langacker (1982) for instance who discusses a theory of linguistics which used to be called Space Grammar and for which this apparently correct fact about human cognition is a central point. Pawley too (see especially, 1980) has argued for allowing syntactic chunks to be stored.

In summary, then, there is nothing in the way that a speaker's own idiosyncratic knowledge of grammar is instantiated that would lead one to believe that word-formation enjoys a separate existence psychologically

from the remainder of grammatical knowledge. Productivity, in particular, does not serve as a criterion for dividing word-formation from the rest of the grammar.

5.2.2 On the Psychological Reality of the Lexical Level.

In the last chapter we examined the question as to whether Lexical Phonology, by positing a lexical level, corresponding roughly to the taxonomic phonemic level, is able to help understand the acquisition of phonology. It was argued that while such a level is surely necessary, it is not clear that the acquisition of the phonology is made simpler in the long run: the lexical level will have to be recovered without benefitting from the advantages LPM has over Taxonomic Phonemics with respect to the theoretical status of that level, since nothing about the lexical phonology of the language in question would be known at that phase.

In this section I examine some other claims, due to Mohanan (1982), that the lexical level is explanatory with respect to "psycholinguistic" data. In particular, I will be looking at speech errors, language games, and speech recognition. I shall suggest that insofar as Mohanan's claims that language processing are correct they are evidence for any theory such posits such a level of representation as the lexical level or word level. I shall point out, however, that there is evidence that language processing accesses (productive) rules at levels more deeply emvbedded than the

lexical level, a fact which is at least consistent with my approach to morphology insofar as while the lexical level or word level is important, it is not sanctified as the output of the lexicon qua word-formation component as it is in LPM. Other levels of representation should thus be accessible to language processing. This is not so clearly true of at least Mohanan's conception of LPM, where the lexicon seems to be viewed as a domain opaque to language processing.

5.2.2.1 Speech Errors.

It has long be argued (c.f. Garrett, 1976) that speech errors provide information about how language is processed. Mohanan argues, on the basis of phonological speech errors, that language production crucially makes use of the lexical level of representation in that such errors always involve the transposition substitution or copying of lexical level phonemes. For instance, one never finds examples like linear set transposed to *l[e]near s[ay]t, but only to l[e]near s[I]t. So it makes no difference that the underlying vowel in the first morpheme of linear is tense (c.f., line), but only that its lexical level form is lax. Alternately, we do not find postlexical processes such as aspiration in English transposed with their segments: s[t]otch [kh]ape is an expected error for scotch tape, never *s[th]otch [k]ape. So, while speech production seems to work off of the lexical level, it is not so late as to apply after post-lexical rules have applied.

I believe that this is essentially correct, although one has to be slightly careful. One does, for instance, see examples of lexical rules applying in speech errors: Misapplications of Ablaut in cases like brang for brought are evidence of this. Since so many verbs ending in ing have this Ablaut pattern it is perhaps not too surprising that a speaker should overextend it in some cases. But this would appear to mean that access is being made by language processing to levels deeper than the lexical level.

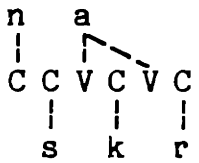
5.2.2.2 Language Games.

Mohanan notes that Secret Languages do not seem to access levels of representation lower than the lexical level. So the ayb Secret Language, which is based on English, must apply after lexical rules, for instance, velar softening: criticism becomes kraybítaybiSaybizaybəm and not *kraybítaybikaybizaybəm. On the other hand, it cannot apply after aspiration: páinter becomes paybéyntaybər and not *p^{h}aybéyntaybər.

Still, McCarthy (in progress) discusses Secret Languages which do access earlier levels of representation. So, in a Bedouin Hijazi Arabic secret language, encoded words are produced by interchanging the consonants of the root. So, a word like dafaʕ might become daʕaf, fadaʕ and so on. However, a word like nkasar in which only the bold-face consonants are part of the root, the /n/ will be left in place and the root consonants interchanged: nsakar, nsarak, nrakas, and so forth. A word with a quadriliteral root

such as tarzam will allow all four consonants to interchange: ratmaz, razmat and so forth. McCarthy argues that this shows that speakers of Bedouin Hijazi have access to the root level of representation. This is encoded by assuming that the scrambling applies only on the root tier (McCarthy, 1979):

(8)



In this representation, /s/, /k/ and /r/ will be interchangeable, but not any of the other segments.

Needless to say, this involves access to a level of representation below the lexical level since we may fairly assume that at the latter level, words are represented only as strings of segments with accompanying prosodic information. Still, it is significant that here too, the language game is making use of a part of the language that is very productive. As McCarthy (1979) argued, Arabic speakers are well aware of the root and pattern nature of the morphophonology of their language, and it is hardly surprising that speech production should have access to this.

5.2.2.3 Speech Recognition.

Finally, Mohanan (p. 145) suggests that it "would not be unreasonable to assume that speech recognition involves the identification of the lexical representation of phonetic strings." On this view, a speech recognition system would not have to concern itself overly much--or indeed at all--with the behavior of lexical rules. This idea has more recently been made use of in the design of a speech recognition system by Church (1983). He states:

> The lexical rules will be used in the parser in an indirect fashion as cues for constituent boundaries. So, for example, we will use the restricted distribution of lax vowels, which is an indirect consequence of lexical rules such as laxing, in order to constrain the possible foot structures.
>
> On the other hand, postlexical rules will play a much more major and direct role for a variety of reasons. First, postlexical rules are easier to apply in the reverse direction than lexical rules because postlexical rules are
>
> (284a) acyclic,
> (284b) free of lexically marked exceptions, and
> (284c) blind to the lexical category and internal morphological structure of the word.
>
> Moreover in addition to the fact that postlexical rules are easier to run in the reverse direction, there are several reasons why they almost have to be run in reverse at run time by the speech recognition device. In particular, postlexical rules are
>
> (285a) the sole source of phonetic allophones, and
> (285b) the only rules that cross word boundaries.

I believe that this is largely correct, though again we must be cautious. For one thing, we can hardly seriously expect to eliminate the need to process lexical rules altogether; as Church suggests, there is

useful information which is to be derived from them. Furthermore, in languages with more complex morphology than English, such as Finnish or Malayalam, such lexical processing becomes almost a necessity since it is not reasonable to expect that all possible words in all of their inflected forms are going to be listed in the lexicon.

What I think there is a good argument for is factoring the task of speech recognition into two parts, the first of which is speech recognition proper, which is concerned only with the discovery of the lexical level of representation, and the second of which is the untying of the lexical phonological rules, a task which has come to be known as morphological recognition; for systems of the latter type see Koskenniemi (1983), Karttunen (1983) and also Sproat (1984).

In this section we have examined the relevance of the lexical level of representation for various psycholinguistic and computational processes. On the whole, Mohanan's point is well taken, though of course, the facts are evidence for any theory which crucially makes use of something equivalent to the lexical level or word level of phonological level of representation, not just LPM. Still, we have to be somewhat careful insofar as it seems as if almost any process, lexical or not, which is perceived by the speakers of a language to be productive, or characteristic of the language, is open for incorporation into secret code languages or

speech errors.

5.3 Some General Conclusions and a Prospectus.

In this dissertation I have argued for a very straightforward point: there is no separate word formation component, what has come to be termed the "Lexicon" in recent years. Rather, I have suggested, the well-formedness of various aspects of word structure is taken care of by principles applying in the various components of the grammar. In Chapters 2 and 3, I showed that the syntactic behavior of a large class of lexical constructions was essentially derivative of principles which are familiar from syntax. In Chapter 4 I suggested that there may not be as great a distinction between lexical and postlexical phonological rules as has been suggested in LPM, and that, in any event, the phonological component insofar as it has a structure similar to that of LPM, should be considered to be a checker of well-formedness of phonological representations, and nothing more. In Chapter 1 I argued that there is a very simple mapping relation between the two levels of representation which I have posited.

What I have not done here is to provide a theory of morphology. What I have done is to provide the rudiments of a theory of morphosyntax and an outline of a theory of morphophonology, and a mapping relation which holds between the two levels. Given what I have said, I could not have provided

a theory of morphology per se, since morphology itself is not a unified entity.

It may be thought that if we take this approach we will have returned to the pre-Remarks days when, for instance, morphosyntax and syntax were not distinct entities[12]. Well, in a sense this is true, since we would no longer have access to a conceivably distinct set of principles applying "in the lexicon." In fact, to my knowledge, nobody has convincingly argued that they are distinct anyway. Taken another way, it is false since the approach I have taken here is lexical in a very important sense: so, in Chapter 2 I argued that there was an abstract affix NOM which attaches to verbs and produces nouns which refer to events; the (syntactic) well-formedness of such nouns would be checked at the syntactic levels of representation. However--and this is the crucial point--we have not made use of powerful transformations in the sense of Lees (1959) for taking whole sentences and transforming them into nominals. Such approaches, or rather their modern equivalents, may be appropriate for some constructions such as gerunds, which appear at D-structure to display sentential properties and at S-structure to display nominal properties (see Baker, 1985b). But our nominalizations are not "derived" in the syntax; a noun

---

12. Note that Aronoff (1976) pointed out that Chomsky's (1970) paper was the primary incentive within Generative Grammar to start looking seriously at morphology insofar as the relationship between destroy and destruction and other such verb-noun pairs was pushed out of the syntax.

like destruction is a noun at all syntactic levels of representation. And this is, after all, all that Chomsky meant by the Lexicalist Hypothesis, namely that the syntax itself does not have transformations which would actually handle the relationship between words like destroy and destruction.

One of the less satisfying aspects of LPM, as discussed in the previous chapter, is the way in which it stipulates the order among lexical processes. In many ways, as I tried to show there in my discussion of English and Malayalam compounding, it would be useful to eschew the Stratum ordering "explanation" for the apparent ordering of lexical processes and try to look more deeply into the matter. I have tried to provide a framework in which this is possible and I feel that much can be learnt from a serious study of morphosyntax, morphophonology and the relationship between them, without making use of the idea of a highly structured lexicon; that structure is surely derivative of more basic principles.

Before closing it would do well to point out that this is by no means the first time that the distinction between morphology and other components of the grammar, especially syntax, has been denied. The idea was of course implicit in Chomsky and Halle (1968), but it predates that. For instance de Saussure (1959) made the following claim (p. 136):

> Functionally, therefore, the lexical and the syntactical may blend. There is basically no distinction between any word that is not a simple, irreducible unit and a phrase, which is a syntactical fact. The arrangement of the subunits of the word

> obeys the same fundamental principles as the arrangement of groups of words in phrases.
>
> In short, although the traditional divisions of grammar [i.e. morphology and syntax: RS] may be useful in practice, they do not correspond to natural distinctions. To build a grammar, we must look for a different and higher principle.

More recently, Langacker (1982: p. 38) has claimed that "natural divisions....do no coincide neatly from one parameter to the next in such a way that we would be justified in carving up the lexicon-morphology-syntax spectrum into discrete components."

It must always be borne in mind that the commitment to the non-existence of a word-formation component does not entail the commitment to the non-existence of words; the second simply does not follow from the first. In fact, it is far from clear that the existence of words has any bearing on the "right theory" of morphology insofar as just saying that there is a separate component of word formation does not go far in explaining what words are. In any event, if we proceed in the fashion I have proposed here and take seriously the idea that morphosyntax and morphophonology should be treated largely separately and as parts of syntax and phonology respectively, then it will, I believe, turn out that much of the structure encoded in recent theories of morphology is reducible to other principles. In this way we can make strides in understanding what is going on in morphology through deriving the lexicon.

Bibliography.

Allen, Margaret. (1978). Morphological Investigations. UConn Ph.D. dissertation.

Anderson, Mona. (1977). "Transformations in NPs." ms. UConn.

_____. (1979). Noun Phrase Structure. UConn Ph.D. dissertation.

Archangeli, Diana. (1984a). "An Overview of the Theory of Lexical Phonology and Morphology." MIT Working Papers in Linguistics. vol. 7, 1-16.

_____. (1984b). Underspecification in Yawelmani Phonology and Morphology. Ph.D. dissertation, MIT.

Aronoff, Mark. (1976). Word Formation in Generative Grammar. Cambridge, MA: MIT Press.

_____, and S.N. Sridhar. (1983). "Morphological Levels in English and Kannada, or Atarizing Reagen." CLS Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax.

Aronson, Howard I. (1982). Georgian: A Reading Grammar. Columbus, OH: Slavica.

Baker, Mark. (1984). "Noun Incorporation in Iroquoian." ms. MIT.

_____. (1985a). "The Mirror Principle and Morphosyntactic Explanation." Linguistic Inquiry. 16(3), 373-416.

_____. (1985b). "Syntactic Affixation and English Gerunds." Paper presented at the 4th West Coast Conference in Formal Linguistics, UCLA.

_____. (1985c). Incorporation: A Theory of Grammatical Function Changing. MIT Ph.D. dissertation.

_____, Kyle Johnson and Ian Roberts. (in prep.). "Explicit Passive Arguments." ms. MIT.

Barss, Andrew. (1984). "Chain Binding." ms. MIT.

Barwise, Jon and John Perry. Situations and Attitudes. Cambridge, MA: MIT Press.

Berko, Jean. (1958). "The Child's Learning of English Morphology." Word 14: 150-77.

Blake, Barry. (1969). The Kalkatungu Language: A Brief Description. Canberra: Australian Institute of Aboriginal Studies.

Brame, Michael. (1984). "Ungrammatical Notes 6: Associativity of Lexical Composition in Syntax and Morphology." Linguistic Analysis 13: 119-126.

Bresnan, Joan. (1982). "Control and Complementation." In Bresnan, Joan, ed.

_____. (ed.) (1982). The Mental Representation of Grammatical Relations. Cambridge, MA: MIT Press.

Booij, Geert. (1983). "Coordination and Reduction in Complex Words: A Case for Prododic Phonology." Vrije Universiteit Working Papers in Linguistics, Amsterdam.

Borer, Hagit. (1984). Parametric Syntax: Case Studies in Semitic and Romance Linguistics. Dordrecht: Foris.

Burzio, Luigi. (1981). Intransitive Verbs and Italian Auxiliaries. MIT Ph.D. dissertation.

Carlyle, Karen. (1983). "Breton Mutation: Linear or Autosegmental?" Paper presented at the Canadian Linguistics Association.

Chomsky, Noam. (1964). Current Issues in Linguistic Theory. The Hague: Mouton.

_____. (1970). "Remarks on Nominalization." in R. Jacobs and P.S. Rosenbaum (eds.) Readings in English Transformational Grammar. Waltham, MA: Ginn.

_____. (1981). Lectures on Government and Binding. Dordrecht: Foris.

_____. (1984). Knowledge of Language. ms. MIT.

_____, and Morris Halle. (1968). The Sound Pattern of English. New York: Harper and Row.

Church, Kenneth. (1983). Phrase-Structure Parsing: A Method for Taking Advantage of Allophonic Constraints. Ph.D. dissertation, MIT Department of Electrical Engineering and Computer Science.

Churma, Donald. (1983). "Jets Fans, Raider Rooters, and the Interaction of Morphosyntactic Processes." CLS Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax.

Clark, Robin. (1984). "Control in NP." Proceedings of the Third West Coast Conference in Formal Linguistics.

Crill, Hildred. (1984). "Celtic Mutations and Post-SPE Phonology". University of North Carolina ms.

Davidson, Donald. (1966). "The Logical Form of Action Sentences." in Davidson, Donald, (1980), Essays on Actions and Events. Oxford: Clarendon Press.

Downing, Pamela. (1977). "On the Creation and Use of English Compound Nouns." Language 53, 810-842.

Fabb, Nigel. (1984). Syntactic Affixation. MIT Ph.D. dissertation.

Fodor, J., M. Garrett, E. Walker, and C.H. Parkes. (1980). "Against Definitions." Cognition 8.

Fowler, H. W. (1926). A Dictionary of Modern English Usage. London: Oxford University Press.

Fromkin, Victoria A. (1973). Speech Errors as Linguistic Evidence. The Hague: Mouton.

Garrett, Merrill F. (1976). "Syntactic Processes in Sentence Production." In R.G. Wales and E.C.T. Walker (Eds.) New Approaches to Language Mechanisms. Amsterdam: North Holland. Pp. 231-255.

Giorgi, Alessandra. (1984). "On C-command and Binding within NPs." ms. Instituto di Psicologia del CNR.

Goodall, Grant. (1984). Parallel Structures in Syntax. UCSD Ph.D. dissertation.

Gordon, Peter. (1984). "Of Mice-Eaters and Men-Eaters: Level Ordering in Young Lexicons." ms. MIT Department of Psychology.

Grimshaw, Jane. (1982). "On the Lexical Representation of Romance Reflexive Clitics." in Bresnan, Joan, ed.

_____, and Ralf-Armin Mester. (1985). "Complex Verb Formation in Eskimo." Natural Language and Linguistic Theory, 3 1-22.

Gruber, Jeffrey. (1965). Studies in Lexical Relations. MIT Ph.D. dissertation.

Guerssel, Mohand. (1983). "A Lexical Approach to Word Formation in English." Linguistic Analysis 12: 183-243.

_____. (1985). "On Phonological Words, Level Ordered Morphology, and Lexical Relatedness." ms. MIT.

Hale, Kenneth, Ann Farmer, David Nash and Jane Simpson. (1981). A Preliminary Dictionary of Lardil. ms, MIT.

Halle, Morris. (1959). The Sound Pattern of Russian. The Hague: Mouton.

_____. (1973). "Prolegomena to a Theory of Word-Formation." Linguistic Inquiry. 4: 3-16.

_____ and K.P. Mohanan. (1985). "Segmental Phonology of Modern English." Linguistic Inquiry. 16: 57-116.

Hammond, Michael. (1984). "Level Ordering, Inflection and the Righthand Head Rule." MIT Working Papers in Linguistics, vol. 7.

Hargus, Sharon. (1985). The Lexical Phonology of Sekani. UC Berkeley Ph.D. dissertation.

Harris, James. (1983). Syllable Structure and Stress in Spanish: A Nonlinear Analysis. Cambridge, MA: MIT Press.

Hayes, Bruce. (1980). A Metrical Theory of Stress Rules. MIT Ph.D. dissertation.

Higginbotham, James. (1985a). "A Note on Phrase Markers." MIT Working Papers in Linguistics volume 6, 87-101.

_____. (1985b). "On Semantics." to appear, Linguistic Inquiry 16(4).

Jaeggli, Osvaldo. (1984). "Passives, Middles and Implicit Arguments," ms., University of Southern California.

Jeanne, LaVerne, Paula Pranka and Kenneth Hale. (1984). "Where's Suppletion?" ms., MIT.

Jespersen, Otto. (1909-49). A Modern English Grammar on Historical Principles. London: George, Allen and Unwin.

Johnson, Kyle. (1985). "Subjects and Theta Theory." ms. MIT.

Karttunen, Lauri. (1983). "KIMMO: A Two Level Morphological Analyzer." Texas Linguistic Forum, 22, Austin, Texas.

Kayne, Richard. (1983). "Unambiguous Paths." in Kayne, Richard, Connectedness and Binary Branching, Dordrecht: Foris.

Kean, Mary-Louise. (1974). The Theory of Markedness in Generative Grammar. MIT Ph.D. dissertation.

Keyser, S. Jay and Tom Roeper. (1984). "On the Middle and Ergative Constructions in English." Linguistic Inquiry 15: 381-416.

Kiparsky, Paul. (1983a). "Lexical Morphology and Morphology." in I.-S. Yang, (ed.) Linguistics in the Morning Calm, 3-91. Seoul: Hansin.

_____. (1983b). "Some Consequences of Lexical Phonology." ms. MIT.

_____. (1983c). "Word Formation and the Lexicon." Mid-America Linguistics Conference.

_____. (1984). "The Lexical Phonology of Icelandic " ms. Xerox Palo Alto Research Center.

Koopman, Hilda. (1984). The Syntax of Verbs. Dordrecht: Foris.

Koskenniemi, Kimmo. (1983). Two-Level Morphology: A General Computational Model for Word-Form Recognition and Production. University of Helsinki Ph.D. dissertation.

Klavans, Judith. (1982). Some Problems in a Theory of Clitics. Bloomington: Indiana University Linguistics Club.

Langacker, Ronald. (1982). "Space Grammar, Analyzability and the English Passive." Language. 58: 22-80.

Lasnik, Howard and Mamoru Saito. (1984). "On the Nature of Proper Government." Linguistic Inquiry 15, 235-289.

Lebeaux, David. (1984). "Nominalizations, Argument Structure and the Organization of the Grammar." ms., UMass Amherst.

Lees, Robert B. (1959). The Grammar of English Nominalizations. MIT Ph.D. dissertation.

Levin, Beth and Malka Rappaport. (1985). "The Formation of Adjectival Passives." Lexicon Project Working Papers #2, MIT Center for Cognitive Science.

Levin, Juliette. (1985). A Metrical Theory of Syllabicity. MIT Ph.D. dissertation.

Lieber, Rochelle. (1980). On the Organization of the Lexicon. MIT Ph.D. Dissertation.

_____. (1983a). "New Developments in Autosegmental Phonology: Consonant Mutation." Proceedings of the Second West Coast Conference in Formal Linguistics.

_____. (1983b). "Argument Linking and Compounds in English." Linguistic Inquiry 14, 251-286.

Lindner, Susan. (1981). A Lexico-Semantic Analysis of English Verb Particle Constructions with OUT and UP. UCSD Ph.D. dissertation.

Lobeck, Anne and Ellen Kaisse. (1984). "On the Domain of Locality Conditions." Proceedings of the third West Coast Conference in Formal Linguistics.

Marantz, Alec. (1981). On the Nature of Grammatical Relations. MIT Ph.D. dissertation.

_____. (1984a). On the Nature of Grammatical Relations. Cambridge, MA: MIT Press.

_____. (1984b). "On the Non-Distinctness of Derivational and Inflectional Morphology." ms. Harvard.

_____. (1984c). "Tagalog Reduplication is Affixation, Too." Paper presented at the LSA, Baltimore.

Marchand, Hans. (1969). The Categories and Types of Present-Day English Word-Formation. 2nd Edition. Munich: Beck.

Mascaró, Joan. (1976). Catalan Phonology and the Phonological Cycle. MIT Ph.D. dissertation.

Massam, Diane. (1983). "The Morphology of Irish Mutation." MIT Working Papers in Linguistics, volume 5.

May, Robert. (1977). The Grammar of Quantification. MIT Ph.D. dissertation.

McCarthy, John. (1979). Formal Problems in Semitic Morphology and Phonology. MIT Ph.D. dissertation.

_____. (in progress.) Formal Morphology.

McCloskey, James, and Kenneth Hale. (1984). "On the Syntax of Person-Number Inflection in Modern Irish." Natural Language and Linguistic Theory. 4: 441-486.

Mchombo, Sam. (1978). A Critical Appraisal of the Place of Derivational Morphology within Transformational Grammar, Considered with Primary Reference to Chichewa and Swahili. University of London, Ph.D. dissertation.

Mohanan, K.P. (1982). Lexical Phonology. MIT Ph.D. dissertation.

Mohanan, K.P. and Tara Mohanan. (1984). "Lexical Phonology of the Consonant System in Malayalam." Linguistic Inquiry 15, 575-602.

Muysken, Pieter. (1981). "Quechua Word Structure." in Frank Heny, (ed.) Binding and Filtering. Cambridge, MA: MIT Press.

Nash, David. (1979). Topics in Warlpiri Grammar. MIT Ph.D. dissertation.

Pawley, Andrew. (1980). "On Meeting a Language that Defies Description by Ordinary Means." Paper presented for Kivung Congress of the Linguistic Society of Papua New Guines, Lae.

Perlmutter, David. (1978). "Impersonal Passives and the Unaccusative Hypothesis." Proceedings of the Berkeley Linguistics Society 4, Berkeley, CA.

Pesetsky, David. (1979). "Russian Morphology and Lexical Theory." ms. MIT.

_____. (1982). Paths and Categories. MIT Ph.D. dissertation.

_____. (1985). "Morphology and Logical Form." Linguistic Inquiry 16, 193-246.

Postal, Paul. (1969). "Anaphoric Islands." Papers from the Fifth Regional Meeting, Chicago Linguistic Society.

Pranka, Paula. (1983). Syntax and Word Formation. MIT Ph.D. dissertation.

Pulleyblank, Douglas. (1983). Tone in Lexical Phonology. MIT Ph.D. dissertation.

Randall, Janet. (1982). Morphological Structure and Language Acquisition. UMass Ph.D. dissertation.

_____. (1984). Morphological Complementation. MIT Working Papers in Linguistics, vol. 7.

Rappaport, Malka. (1983). "On the Nature of Derived Nominals." in Levin, Lori (et. al.) Papers in Lexical-Functional Grammar, Bloomington: Indiana University Linguistics Club.

_____, and Beth Levin (1985). "A Case Study in Lexical Analysis: the Spray/Load Alternation." unpublished ms., MIT.

Reinhart, Tanya. (1976). The Syntactic Domain of Anaphora. MIT Ph.D. dissertation.

Roeper, Thomas. (1984). "Implicit Arguments." ms., UMass Amherst.

_____, and Muffy Siegel. (1978). "A Lexical Transformation for Verbal Compounds." Linguistic Inquiry 9, 199-260.

Rothstein, Susan. (1983). The Syntactic Forms of Predication. MIT Ph.D. dissertation.

Sadock, J. (1980). "Noun-Incorporation in Greenlandic: a Case of Syntactic Word Formation." Language 56, 300-319.

de Saussure, Ferdinand (1959). Course in General Linguistics. trans. by Wade Baskin. New York: McGraw Hill.

Selkirk, Elizabeth O. (1980). "The Role of Prosodic Categories in English Word Stres," Linguistic Inquiry 11, 563-605.

_____. (1982). The Syntax of Words. Cambridge, MA: MIT Press.

_____. (1984). Phonology and Syntax: The Relation between Sound and Structure. Cambridge, MA: MIT Press.

Shipley, William. (1964). Maidu Grammar. University of California Publications in Linguistics, vol. 41. Berkeley: University of California Press.

Siegel, Dorothy. (1974). Topics in English Morphology. MIT Ph.D. dissertation.

Simpson, Jane. (1983). Aspects of Warlpiri Morphology and Syntax. MIT Ph.D. dissertation.

Speas, Margaret. (1984). "Navajo Prefixes and Word Structure Typology." MIT Working Papers in Linguistics, vol. 7.

Sproat, Richard. (1982). "Redundancy Rules and Welsh Mutation." ms. MIT.

_____. (1984a). "On Bracketing Paradoxes." MIT Working Papers in Linguistics, vol. 7.

_____. (1984b). "Parsing Morphology." Technical Memorandum, AT&T Bell Laboratories, Murray Hill, NJ.

_____. (1985a). "Welsh Syntax and VSO Structure." Natural Language and Linguistic Theory 3: 173-216.

_____. (1985b). "A Note on Bracketing Paradoxes." MIT Working Papers in Linguistics, vol. 6.

Stemberger, Joseph. (1982). The Lexicon in a Model of Language Production. UCSD Ph.D. dissertation.

Stowell, Timothy. (1981). Origins of Phrase Structure. MIT Ph.D. dissertation.

Strauss, Steven L. (1982a). "On 'Relatedness Paradoxes' and Related Paradoxes." Linguistic Inquiry 13, 695-700.

_____. (1982b). Lexicalist Phonology of English and German. Dordrecht: Foris.

Thomas-Flinders, Tracy. (1981). "Initial Lenition in Celtic: Evidence for Inflection as a Phonological Process." in Thomas-Flinders, T. (ed.) Inflectional Morphology: Introduction to the Extended Word-and-Paradigm Theory. University of California, Los Angeles.

_____. (1983). Morphological Structures. University of California, Los Angeles Ph.D. dissertation.

Thomason, Richmond. (1985). "Some Issues Concerning the Interpretation of Derived and Gerundive Nominals." Linguistics and Philosophy, 8: 73-80.

Travis, Lisa. (1984). Parameters and Effects of Word Order Variation. MIT Ph.D. dissertation.

Watkins, M. (1937), A Grammar of Chichewa. Language Dissertations, vol. 24.

Whitney, William Dwight. (1889). Sanskrit Grammar. Cambridge, MA: Harvard University Press.

Williams, Edwin. (1980). "Predication." Linguistic Inquiry 1, 81-114.

_____. (1981a). "On the Notions 'Lexically Related' and 'Head of a Word'." Linguistic Inquiry 12, 245-274.

_____. (1981b). "Argument Structure and Morphology." The Linguistic Review. 1, 81-114.

_____. (1982). "The NP Cycle." Linguistic Inquiry 13.

_____. (1984). "Coanalysis." Work in progress, UMass Amherst.

Yip, Moira. (1978). "The Integrity of Lexical Nodes." ms. MIT.