INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeob Road
Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
RODMAN, Robert David, 1940-
THE STUDY OF FUZZY ISLANDS WITHIN THE
FRAMEWORK OF TRANSFORMATIONAL GENERATIVE
GRAMMAR.

University of California, Los Angeles,
Ph.D., 1973
Language and Literature, linguistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The Study of Fuzzy Islands
within the Framework of Transformational Generative Grammar

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Linguistics
by

Robert David Rodman

1973
The dissertation of Robert David Rodman is approved, and it is acceptable in quality for publication on microfilm.

George F. Bodell, Committee Chairman

Victoria A. Fromkin

David S. Gillman

University of California, Los Angeles

1973
TO MY PARENTS

NATHAN AND SONIA RODMAN
ACKNOWLEDGMENT

I wish to express gratitude to my Thai informant Pimpaporn Suwattigul for patiently understanding my exasperating penetrations of her grammatical intuitions. My thanks also to Sandra Thompson and Georgette Silva, and their informants, for helping me obtain data in Mandarin Chinese and Korean.

Thanks also to Barbara Hall Partee for "turning me on" to Zadeh and his work, and for criticising some of my early ideas. I am especially grateful to my friend Victoria Fromkin who carefully pointed out many misapprehensions during the early stages of writing. Professor Fromkin's continual encouragement and confidence were a source of strength to me whenever the going got tough. My advisor, George Bedell, cogently and meticulously criticised an early version of chapter two, and assisted me with the entire Dissertation. I owe him many thanks.

I am also grateful to the Department of Linguistics, University of California, Los Angeles, for providing four years of financial support in the form of a Chancellor's Teaching Fellowship. I am indebted as well to the National Science Foundation, and to principal co-investigators Paul Schachter and George Bedell, for providing employment in the summers of 1971 and 1972 under the research grant GB-3204. And my thanks to Theodora Graham for typing the final draft with her usual efficacy.
# TABLE OF CONTENTS

| Acknowledgement                                      | iv   |
| Vita                                                | vi   |
| Abstract of the Dissertation                        | viii |
| 1 Introduction                                      | 1    |
| Footnotes                                           | 13   |
| 2 The notion "Island"                               | 14   |
| 2.1 Historical Background                           | 14   |
| 2.2 Inadequacies of Ross's Island Constraint        | 29   |
| 2.3 The Foreground/Background Constraint            | 46   |
| Footnotes                                           | 89   |
| 3 Islands and the Island Constraint                 | 93   |
| 3.1 Strengthening and Weakening Factors             | 93   |
| 3.2 The Island Constraint                           | 138  |
| 3.3 NP Islands                                      | 230  |
| Footnotes                                           | 256  |
| 4 Implications of this work and other fuzzy aspects of grammar | 265  |
| Footnotes                                           | 275  |
| Bibliography                                        | 276  |
| Appendix                                            | 278  |
VITA

1961  BA  Mathematics  UCLA
1965  MA  Mathematics  UCLA
1971  Teaching Assistant and Research Assistant
      (Summer), Department of Linguistics, UCLA
1971  MA  Linguistics  UCLA
1972  Teaching Associate and Research Associate
      (Summer), Department of Linguistics, UCLA
1972  C. Phil. Linguistics  UCLA

AWARDS

1969 - 1973  Chancellor’s Teaching Fellowship

PUBLICATIONS

"A critical look at some movement constraints," in Critiques of
Syntactic Studies I, UCLA Papers in Syntax No. 1, edited by
Paul Schachter and George Bedell. 1971.

"On the stranding of prepositions in English," in Explorations in
Syntactic Theory, UCLA Papers in Syntax No. 2, edited by George
Bedell. 1972.

"Movement constraints on transformations that do not move elements," in
Studies in East Asian Syntax, UCLA Papers in Syntax No. 3,

"The proper treatment of relative clauses in a Montague Grammar," in
Papers in Montague Grammar, UCLA Occasional Papers in
PUBLICATIONS EDITED

Papers in Montague Grammar, UCLA Occasional Papers in Linguistics

TEXTBOOK

Introduction to Language. To be published in 1973 by Holt, Rinchart and Winston. Co-author: Dr. Victoria Fromkin
ABSTRACT OF THE DISSERTATION

The Study of Fuzzy Islands

within the Framework of Transformational Generative Grammar

by

Robert David Rodman

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 1973

Professor George D. Bodell, Chairman

In the past several years linguists have become increasingly interested in non-discrete, or fuzzy linguistic phenomena. The law of the excluded middle, once thought to hold pervasively in a transformational grammar, appears not to hold in many instances. For example certain rules may not simply be applicable or non-applicable, but applicable to a degree. Certain constraints on the grammar may not be operable or inoperable, but operable to a degree. Surface structures are not grammatical or ungrammatical, but grammatical to a degree (or ungrammatical to a degree). Certain grammatical configurations are not necessarily islands or non-islands, but islands to a degree. The latter statement is the subject of the Dissertation.

My study of English and Thai principally, and Korean, Mandarin Chinese, Japanese and Kannada marginally, have led me to put forth certain specific proposals about islands: All non-root S's and NP's are fuzzy islands. They are acted on by a variety of island-strengthening and island-weakening processes.
both syntactic and semantic, in such a way that a wide range of island strengths are observed, from very weak (and non-islandlike) to very strong (such as a non-restrictive relative clause island). The strength of an island is a function of the entire derivation of the structure containing the island.

To capture the notion of "island to a degree" or "fuzzy island" I have drawn on L. Zadeh's theory of fuzzy sets. An island is described as a fuzzy set of nodes of the phrase structure tree. The strength of a node within its island is proportionate to the degree of membership of the node in this fuzzy set. By this means we are able to describe strong islands as being associated with fuzzy sets containing elements with high degrees of membership. We also can capture the fact that some nodes have stronger membership in an island than other nodes.

The appropriate island constraint is not, I propose, a constraint on individual rules, but a well-formedness constraint on grammatical structures. The island constraint considered in the text is simple: Islands may not overlap. To the degree that islands overlap surface structures will be ungrammatical.

Several advantages of incorporating these proposals into the grammar are the elimination of the distinction between chopping rules and copying rules insofar as islands are concerned, the elimination of the Left Branch Condition on piedpiping, the elimination of the Sentential Subject Constraint, and the possible elimination of the need for the notion of Bounding. Moreover, by viewing islands as quanta of linguistic structure which speakers
feel compelled to maintain whole and distinct, we see that the
edict against overlapping islands has a psychological reality
that island constraints such as the Complex NP Constraint lack.

The constraint against overlapping islands is a
candidate for a universal principle of language. The determination
of what comprises an island, and what factors act to strengthen
and weaken islands, are likely to be language specific phenomena.
CHAPTER ONE

One of the most perplexing situations a beginning student of linguistics faces comes about when he takes literally the following statement by Chomsky and tries to implement it in his native language.

The fundamental aim in the linguistic analysis of a language \( L \) is to separate the grammatical sequences which are the sentences of \( L \), from the ungrammatical sequences which are not sentences of \( L \).  [Syntactic Structures, Chapter 2]

Almost before he's memorized "colorless green ideas sleep furiously" the student is confronted with a puzzling array of data which he simply cannot bifurcate into two classes on the basis of grammaticality (or acceptability). Moreover, he discovers, everybody else has the same problem. To compound matters, when he forces himself to decide on the status of a particular sequence, he not only finds himself frequently in disagreement with his colleagues and teachers, he finds himself, the next day or so, in disagreement with judgments he made on the previous day. In (1.1) a smattering of data likely to confuse is given.¹

(1.1) We got ourselves misunderstood.

John saw the letter opened which was from Mary.

She is a woman who Frank is shorter than.

We dislike to be so coy.

What Caesar refused was to submit to the senate.

1
Lees pinpoints the problem in his Review of Syntactic Structures (1957) where he states:

Another internal technical problem which awaits exact solution is entailed by the establishment of a grammatical theory upon the notion of grammaticalness of sentences. It is necessary to establish some scale of grammaticalness along which every utterance will lie and which will correspond well to our intuitive feelings about how sentences are construed, and which, furthermore, will be automatically derivable from the general theory of language or from the grammar of each particular language. This is to say that a grammar must explicate our notion that certain structures, while very bizarre, are nevertheless not completely excluded by the pattern of the language, and that certain ones are less excluded than others and perhaps even by a specifiable degree.

If sentences have varying degrees of grammaticalness then it is to be expected that a large class of sequences will be difficult to classify as either grammatical or ungrammatical. Moreover, much of the variation observed among speakers can be attributed to the fact that in discretizing what is inherently non-discrete, insignificant idiosyncratic factors can cause a judgment to swing one way or the other. An analogous situation is a one hundred ton boulder balanced precariously on end. A gust of wind or a child could cause it to topple in any direction.

Idiolectal variation, in many cases, is likely to revolve around the degree of grammaticality of certain sequences. In fact, this is precisely the place where one ought to begin to find such variation since it is often the gray zone of grammaticality that gives rise to differing
intuitions of native speakers. Thus when speakers are divided on the status of questionably grammatical sentences we are tempted to look for idiolectal differences. When speakers are sharply divided on a sentence that one group finds unquestionably grammatical, but the other group finds unquestionably ungrammatical, then we may want to look for dialectal differences.

The same kind of situation obtains when judgments of meaning are attempted. Speakers cannot always decide on the precise meaning of a sentence, and there is an absence of uniformity among speakers.

Lakoff (1973) notes these facts and observes:

Almost every syntactic or semantic phenomenon has a shadowy area in which speakers become unclear with respect to judgments about meaning and well-formedness. For the past few years, this phenomena has been studied intensively by Ross. His general results are as follows:

(i) Rules of grammar do not simply apply or fail to apply; rather they apply to a degree.

(ii) Grammatical elements are not simply members or non-members of grammatical categories; rather they are members to a degree.

(iii) Grammatical constructions are not simply islands or nonislands; rather they may be islands to a degree.

(iv) Grammatical constructions are not simply environments or nonenvironments for rules; rather they may be environments to a degree.

(v) Grammatical phenomena form hierarchies which are largely constant from speaker to speaker, and in many cases, from language to language.

(vi) Different speakers (and different languages) will have different acceptability thresholds along these hierarchies.

Clearly there is research fuel for fifty dissertations (if not five hundred) and some narrowing down must be made.

I have chosen to concentrate my efforts on (iii).
In standard theories a construct is either an island or it is not. If it is, then certain constraints apply, or do not apply, with the end result being a grammatical sequence, or blockage. But such a theory is inherently unable to tell us that in (1.2), despite the fact that each string violates the Complex NP Constraint, the topmost sequences are but marginally ungrammatical, and entries in the list get successively worse as you go from top to bottom.

(1.2) a. 
   b. 
   c. It was Sally who he invented{ the }story that John ran over.
   
   d. 
   e. his own

All the strings in (1.3) are also violations of the Complex NP Constraint, and they too form a hierarchy of grammaticality. But note that the sentences of (1.3) are worse, respectively, than those of (1.2).

(1.3) a. 
   b. 
   c. It was John who he invented{ the }story that ran Sally over.
   
   d. 
   e. his own

These facts are in general agreed upon by native speakers of English and must therefore be represented in a grammar of the language.
Needless to say there are many situations typified by (1.2) and (1.3), all of which point to our abandoning the notion of discrete syntactic island in favor of a more adequate representation.

The changes to the theory of grammar that (i)-(vi) imply are far-reaching. It is not clear that Classical Transformational Grammar is flexible enough to accommodate such changes. Transformational grammarians will have to accept the fact that grammaticality is a function of whole derivations, and that therefore each derivational step may influence the status of the final output. They will also have to accept the additional theoretical machinery necessary to accommodate the various non-discrete notions that will permeate the grammar. In the long run, it is to be hoped, a non-discrete grammar will be "simpler" by virtue of the fact that many of the special "conditions" that presently infect transformational grammars will no longer be necessary. They will have yielded to broader, though non-discrete generalizations.

An example of this is Ross's Internal Island Sentential Noun Phrase Constraint, which in its discrete formulation looks something like this:

Grammatical sentences containing an internal NP which exhaustively dominates an S are unacceptable, unless the main verb of that S is a gerund. [Ross (1967:138)]
In its non-discrete formulation, the same constraint can be expressed without the exception clause for gerunds. (See section 3.2.2.7 for details.)

Grammatical sentences containing an internal NP which exhaustively dominates an S are less acceptable, the nounier the complement.

The basic formal device that I will appeal to is that of the fuzzy set.

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. [Zadeh (1965) "Fuzzy Sets," Information and Control 8, 338-353]

Normally we conceive of a set as having some collection of objects, drawn from a universe of discourse, as members; all other objects in the universe of discourse are non-members. The set of positive integers, of letters on this page, or of people that subscribe to Language are normal sets. Given a universe of objects it is a straight-forward matter to determine membership status for each object with respect to any of the above sets.

Suppose, however, that we choose to define a set of objects by a somewhat more subjective criterion. Consider the set of green objects, where the universe of discourse consists of visible entities. Some objects will be of a hue such that their greenness is uncompromised (admittedly a subjective judgment). These objects will have a membership coefficient of 1.0. Other objects will appear "less green," either leaning toward blue, toward yellow, or perhaps toward
a lesser degree of saturation (i.e., less saturation means more light of various non-green wavelengths mixed in). Their membership coefficient in the fuzzy set of green objects will be less than 1.0.

Other fuzzy sets can easily be imagined. The set of middle-aged persons, of good chess players, of handsome elephants, of tasty ham sandwiches, of bright stars, of bald men, of tall women, of well-behaved children, and so on, are fuzzy sets. Notions which are inherently indiscretizable are best represented by means of fuzzy sets.

Formally, then, we can start by declaring that:

(1.4) A language is a fuzzy set of sentences. The coefficient of membership of a string in the fuzzy set of sentences of a language is dependent on the derivations of that string by the grammar.

Note that we are suggesting that every string generated by the grammar is accompanied by a membership coefficient of that string in the fuzzy set of sentences. We can view this coefficient as a "coefficient of grammaticalness." It is beyond our intuitions to fathom exactly the coefficient of grammaticalness of semi-grammatical sentences, just as it is presently beyond the scope of any theory of grammar I can conceive of to produce these coefficients. This difficulty needn't detain us. What is important is the relative grammaticality of sequences, relative degree of islandness of constructs, etc. Given two semi-grammatical strings we can
often determine their relative grammaticality, although it may be quite difficult to state the coefficient of grammaticality in the individual cases. The grammar we write must capture these ordering relationships. Our concern in this work, then, will be more with relative degree than with absolute magnitude.

We expect speakers to differ on their intuitions about the absolute grammaticality of strings and view this variation as idiolectal for the most part. If speakers differ quite definitely on the grammaticality of two strings relative to each other, we would expect the difference between their grammars to be characteristic of dialectal differences.

That we are concerned with relative degrees of grammaticality, islandness, etc., is not a retreat from my belief that linguistic competence is in many critical ways fuzzy. It is an accommodation to the practical fact that our intuitions have only a limited ability to measure our mental processes. Our internal grammar is, I believe, capable of generating sentences of any degree of grammaticality. Our intuitions, as Lakoff (1973) suggests, are barely capable of fathoming more than five degrees of grammaticality: Good, pretty good, in between, pretty bad, definitely out. However, even within a single category, say the category "pretty good," our intuitions are capable of finer distinctions when it comes to comparing two sequences. Of course at some point, which varies among individuals, idiosyncratic factors make

8
their weight felt and finer distinctions simply become impossible to make. There is an uncertainty principle at work. Our mental processes make finer distinctions than we are consciously aware of, and when we plumb our intuitions and try to get a feeling for the grammaticality of a sentence, that grammaticality may seem to change or, as I often experience, may seem to blur over. It is a little like trying to "see" an electron by focusing a beam of photons (light) on it. The measuring process itself affects the object to be measured.

The notion "island" has never been formalized, but in light of the formal definition of a tree (discussed in the Appendix) it is clear that an island is a certain set of nodes. The notion of "island to a degree" or fuzzy island, is captured by defining the set of nodes that correspond to an island to be a fuzzy set.

Except for changes to the theory directly concerning islands, I have attempted to work within the framework of classical transformational grammar. The reasons are twofold. First, it is well beyond the scope of a doctoral thesis to found a new theory of fuzzy grammar, into which my theory of islands would fit. Second, by showing that my theory of islands and island constraints is (more or less) compatible with an established framework that has, after all, accounted for a vast number of linguistically significant facts, I gain a credibility that might be lacking were the work done within an entirely new and unseasoned theoretical framework.
I wish to anticipate the following criticism: Grammaticality is binary, but there are lots of fuzzy performance factors that lead to non-binary acceptability. In opposition to this view it is my belief that the degree of grammaticalness of a string is one of the factors that contributes to the degree of acceptability—and one of the most important ones.

If grammaticality is binary, then in the absence of performance factors, acceptability must also be binary. But this clearly isn't so, as (1.2) and (1.3) attest, unless one is willing to claim that the choice of determiner is a performance matter. Consider in this vein (1.5):

(1.5) He said (that) He complained ?(that) the music was too loud.
                He hinted ?*(that)

The presence or absence of that affects the acceptability of (1.5bc). This doesn't appear to be a performance factor.

As another example, consider the rule of Particle Movement. In (1.6) as you go from (a) to (d), the sequences become less acceptable.

(1.6) a. I called the man who you know up.

b. I called the nervous, flighty man who your friends know up.

c. I called the man who gave my sheepdog puppy several nice juicy babies to chew on up.

d. I called the man that John mentioned in his letter and who Mary told me was an accomplice of the new Supreme Court Justice G. Gordon Liddy up.
In this case it is clear that the length of the material intervening between the verb and its particle plays a role in the acceptability judgments. It is likely that this is related to short term memory capacity which is relevant to linguistic processing and as such is a (fuzzy) performance factor.

Contrast this with (1.7):

(1.7) a. He looked the number of the station up.
    b. He put their writing a letter down.
    c. He looked what they bought over.
    d. He wrote that John was stupid down.

Again there is decreasing acceptability, but in this case it can be traced to the type of syntactic structure over which the particle moves. These syntactic structures play a pervasive role in the competence grammar, and therefore the decreasing acceptability in (1.7) is in great part a function of decreasing grammaticality. Performance factors in this case, if any, play a minimal role.

We will see many examples of semi-acceptable sentences in which the source of semi-acceptability can be traced to factors that are normally considered a part of the competence of speakers, such as choice of complementizer, choice of determiner, presence of presupposition, whether a rule has operated or not, etc.

Thus I am claiming that linguistic competence is inherently fuzzy. This should not be surprising in view of
the fact that much human thinking and interaction with the world involves classes of entities without sharply defined boundaries, in which the transition from membership to non-membership forms a continuum. As Zadeh states:

The ability of a human brain, weighing only a few hundred grams, to manipulate complicated fuzzy concepts and act on multidimensional fuzzy sensory inputs endows it with a capability to solve rather easily a wide variety of problems which, if formulated in precise quantitative terms, would exceed the computing power of the most powerful, the most sophisticated digital computer in existence. The explanation for this apparent paradox is that, in many instances, the solution to a problem need not be exact, so that a considerable measure of fuzziness in its formulation and results may be tolerable. The human brain is designed to take advantage of this tolerance for imprecision." [Zadeh (1971) "Fuzzy Languages and Their Relation to Human and Machines Intelligence" p. 3]

We assume, then, for the rest of this work, that the linguistic competence of native speakers of all languages is fuzzy in the sense defined in this chapter. We focus our attention now on one aspect of the competence model—syntactic islands.
Footnotes

Chapter One

1Taken from N. J. Spencer's handout at the 1972 Winter LSA meeting. She, in turn, collected them from the linguistic literature. These sentences, and many more, were shown to a large number of naive and non-naive (linguists) informants who were asked to make judgments of grammaticality. The sentences I chose for (1.1) are ones in which the groups divided nearly 50-50.

2Actually, if my proposal is accepted that the node dominating complements is a fuzzy node, then this constraint can be even more simply stated. This proposal is briefly discussed in the Appendix.

3Expressed to me by Barbara Partee in a personal communication.
CHAPTER TWO

2.0

2.1 The seed of the notion island, and of future island constraints, was sown by Chomsky when, in a paper entitled "The Logical Basis of Linguistic Theory," he formulated the so-called A-over-A principle:

"...if the phrase of X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW)." (Chomsky, 1962, 931)

The inadequacies, as well as the strengths, of this principle are well discussed in the literature (cf. Chomsky (1964), (1968), and (1972); Ross (1967)). Ross undertook the task of delimiting in a precise way the various inadequacies of the A-over-A principle, and moreover, set himself the goal of replacing this principle by other principles that would explain all the phenomena that it accounted for, while at the same time avoiding its inadequacies.

The history of this work is, of course, set out by Ross. It is my intention to pass over the arguments against the A-over-A principle and to consider directly two of the (several) principles that were intended to replace it, namely the Complex NP Constraint and the Coordinate Structure Constraint.
Because of sentences like (2.1) the rules WH-Q and WH-REL must be formulated with an essential variable in their structural description:

(2.1) a. Alex discovered he had no power to control his desire for sex and violence.

b. What did Alex discover he had no power to control his desire for?

c. The girl who Alex discovered he had no power to control his desire for went berserk in a clock factory.

But that being the case, what is to block the clearly ungrammatical sentences (2.2bd)?

(2.2) a. Filbert promised to listen to the claim that the Vikings explored the moon.

b. *What did Filbert promise to listen to the claim that the Vikings explored?

c. Otis managed to convince his father that the man that hired his girl friend ran a pizza house for profit.

d. *The girl friend who Otis managed to convince his father that the man that hired ran a pizza house for profit once turned down a movie contract.

Ross considered this problem and, observing similar behavior in other languages, he proposed that the following constraint be incorporated into the theory of grammar (making it a "linguistic universal").

(2.3) COMPLEX NOUN PHRASE CONSTRAINT (Ross 1967:70): No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

In (2.2a), $\text{NP}_{\text{NP}}^{\text{NP}}$ is a noun phrase with the lexical head noun $\text{NP}$.
the claim. The CNPC prevents the moon from being questioned by a preposing operation. Similarly, in (2.2c) the man that hired his girl friend is a relative construction, which is a "noun phrase with a lexical head noun" par excellence. The counterpart to the NP node that dominates the string his girl friend in (2.2c) cannot be moved out of the larger noun phrase the man that hired WH-someone to form a relative clause, as in (2.2d).

Still bearing in mind that the WH-preposing rules must be formulated to move elements over an essential variable, we must somehow account for the ungrammatical sentences in (2.4bd).

(2.4) a. Alex loved sex and violence.

b. *What does Alex love sex and?

c. Alex loved sex and Georgie loved violence.

d. *What did Alex love and Georgie loved violence?

Ross had an answer for this case too. He proposed that (2.5) also belong to the theory of grammar.

(2.5) COORDINATE STRUCTURE CONSTRAINT (Ross 1967:89):
In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

It is clear how (2.5) acts to block the generation of the ungrammatical sentences of (2.4).

Since the CNPC and the CSC are supposed to be universal principles of language, we expect them to hold in Thai. We must try to find chopping rules in Thai, i.e. rules which permute elements over an essential variable, because those
are the only kind of rules that are governed by the constraints (Ross 1967:236).

The process of Object Preposing in Thai moves elements to the head of the topmost S, as shown in (2.6):

(2.6) a. khru: dỳ:m nom kâe:w ní:
     teacher drink milk glass of this

     The teacher drank this glass of milk.

b. nom kâe:w ní: khru: dỳ:m

     This glass of milk the teacher drank.

It is quite clear that (2.6b) is derived from (2.6a)—Thai is an SVO language. That an element can be preposed from indefinitely far away (hence chopped) is suggested by (2.7):

(2.7) a. khâw chỳa wâ: khru: dỳ:m
     he believe complementizer teacher drink

     nom kâe:w ní:
     milk glass of this

     He believes the teacher drank this glass of milk.

b. nom kâe:w ní: khâw chỳa wâ: khru: dỳ:m

     This glass of milk he believes the teacher drank.

To test the CMPC we try to prepose the direct object out of the complex NP shown in (2.8). The ungrammaticalness of (2.9) indicates that the constraint holds.

(2.8) phôm châ:b "NP[khru:] S[thì: dỳ:m nom
     I like COMP teacher drink milk

     kâe:w ní:]"
     glass of this

     I like the teacher who drank this glass of milk.

(2.9) nom kâe:w ní: phôm châ:b khru: thì: dỳ:m

17
In a similar way the CSC can be shown to hold in Thai. It is impossible to prepose either direct object in (2.10), as the data in (2.11) show.

(2.10) khruː dỳːm nom kæːw níː laē chaː teacher drink milk glass of this and tea thuːaj níː cup of this

The teacher drank this glass of milk and this cup of tea.

(2.11) a. *nom kæːw níː khruː dỳːm laē chaː thuːaj níː

b. *chaː thuːaj níː khruː dỳːm nom kæːw níː

We might note in passing that movement within a con-

junct can take place, a result consistent with the CSC.

(2.12) khruː dỳːm nom kæːw níː laē nágríān teacher drink milk glass of this and student dỳːm chaː thuːaj níː drink tea cup of this

The teacher drank this glass of milk and the student drank this cup of tea.

(2.13) nom kæːw níː khruː dỳːm laē nágríān milk glass of this teacher drink and student dỳːm chaː thuːaj níː drink tea cup of this

This glass of milk the teacher drank and the student drank this cup of tea.

Another rule in Thai---call it the rule of Focus, relates

(2.14a) to (2.14b).

(2.14) a. khāːw kʰāːj phâː he sell cloth

He sells cloth.
b. Ḝ: phaː nānḷəe thiː khāw kʰaːj
   particle cloth FOCUS COMP he sell

   It's cloth that he sells.

The NP phaː appears to be reordered into focus position out of a constituent structure that, except for a deep structure element of emphasis or focus, underlies (2.14a). Example (2.15) suggests that the focused element can originate indefinitely far down the tree, suggesting that the rule of Focus is a chopping rule.

(2.15) Ḥ: nāngsyː nānḷəe thiː khāw phίː suːd
   particle book FOCUS COMP he prove
   waː thəː syː
   COMP she buy

   It's a book that he proved that she bought.

The data in (2.16) show that the CNPC and the CSC are obeyed.

(2.16) a. *Ḥ: phaː nānḷəe thiː phuːchaːj thiː
   part. cloth FOCUS COMP man COMP
   kʰaːj suːŋ
   sell tall

   *It's cloth that the man who sells is tall.

b. *Ḥ: phaː nānḷəe thiː khāw kʰaːj
   part. cloth FOCUS COMP he sell
   lāe phonlāmaːj
   and fruit

   *It's cloth that he sells and fruit.

As in so many of the world's languages (e.g. Mandarin Chinese, Japanese, Korean), relative clause formation in Thai is accomplished by deletion of the relativized nominal in the embedded clause under identity with a nominal in the
matrix clause. Significantly, no elements are reordered. In (2.17), the circled NP is deleted from the underlying structure to give the surface utterance. The element \( \hat{\text{thi}}: \) is a complementizer.

\[(2.17)\] \( \text{phôm} \ ch\hat{\text{i}}:b \ d\hat{\text{eg}} \ \hat{\text{thi}}: \ k\hat{\text{amlang}} \ k\hat{\text{hi}}: \ m\hat{\text{á}}: \)  
\( \text{I like} \) \( \text{boy} \) \( \text{COMP} \) \( \text{PROG} \) \( \text{ride} \) \( \text{horse} \)  
\( \text{tua} \ nán \)  
\( \text{classifier} \) \( \text{that} \)

I like the boy who is riding that horse.

Relative Clause Formation must be formulated with an essential variable in its structural description because elements indefinitely far down a tree can be relativized, as suggested by (2.18).

\[(2.18)\] \( \text{phû:cha:j} \ \hat{\text{thi}}: \ \text{phôm} \ k\hat{\text{hi}}: \ w\hat{\text{á}}: \ k\hat{\text{hu} n} \ ch\hat{\text{ya}} \)  
\( \text{man} \) \( \text{COMP} \) \( \text{I think} \) \( \text{COMP} \) \( \text{you believe} \)  
\( w\hat{\text{á}}: \ k\hat{\text{hru}:} \ h\hat{\text{én paj} \ \hat{\text{thi}}: \ \text{tàl\hat{\text{a}:d lae:w} \)  
\( \text{COMP} \) \( \text{teacher} \) \( \text{see} \) \( \text{go to} \) \( \text{market already} \)

The man whom I thought you believed the teacher saw has already gone to the market.

20
However, Thai does not permit elements which themselves belong to relative clauses to be relativized, as we see in (2.19).

(2.19) a. phû:cha:j thi: mae:w hên chô:b mà: man COMP cat see like dog

   tua  nán classifier that

The man who the cat sees likes that dog.

b. *mae:w thi: phû:cha:j thi: hên chô:b cat COMP man COMP see like

   mà: tua  nán  dam
dog classifier that black

*The cat that the man who saw likes that dog is black.

Nor can any element in a coordinate structure be relativized, as seen in (2.20) and (2.21).

(2.20) a. mae:w lâe mà: dû:m nom cat and dog drink milk

The cat and the dog drank milk.

b. *mae:w thi: lâe mà: dû:m nom pen cat COMP and dog drink milk be

   khô:ng chan poss. I

*The cat that and the dog drank milk is mine.

(2.21) a. mae:w dû:m nom lâe phû:cha:j kin kha:w cat drink milk and man eat rice

The cat drank milk and the man ate rice.

b. *kha:w thi: mae:w dû:m nom lâe phû:ch:j rice COMP cat drink milk and man

   kin pen khô:ng chan
cat be poss. I

*The rice that the cat drank milk and the man ate is mine.
In addition to the CNPC and the CSC, an adequate grammar of Thai will need a further constraint to the effect that elements in relative clauses and coordinate structures may not be relativized. But we would then find the grammar cluttered with constraints of relatively low functional use. It is unlikely that such a grammar would reflect the intrinsic knowledge the Thai speaker has about his language. It would be desirable to eliminate one or more of these constraints in favor of a broader, more general principle.

Analogous results are obtainable from English. I'll only give a single example, but it is first necessary to digress somewhat in order to discuss the formation of restrictive relative clauses in English. Most linguists are agreed that the morpheme that, when it introduces a relative clause, is not a true relative pronoun, but rather the same complementizing particle that introduces other kinds of subordinate clauses. Under such an analysis the formation of relative clauses takes place in several steps. The first (obligatory) step would be WH-pronominalization.

i. man (that I swore at the man) \( \rightarrow \) man (that I swore at WH-him)

ii. man (that the man swore at me) \( \rightarrow \) man (that WH-he swore at me)

iii. man (that the man's son swore at me) \( \rightarrow \) man (that WH-his son swore at me)

At this point in the derivation an optional rule of WH-pronoun deletion is encountered. If it applies to (i) or (ii), the result will be, ignoring other circumstances for the moment, the strings (i') and (ii').

22
1'. man that I swore at
2'. man that swore at me
This rule may not apply to (iii) for very general reasons. The element WH-his may not be deleted because of an independently motivated constraint on the grammar that prevents left-branching elements inside noun phrases from being reordered or deleted (see discussion of the Left Branch Condition in section 3.3). Furthermore, the substring WH-his son may not be deleted because of restrictions on the recoverability of deleted material (i.e. son is irrecoverable).

The next relevant rule in the sequence is the obligatory rule of WH-pronoun preposing. If the optional rule of WH-pronoun deletion is not carried out, then the preposing rule will give:

1''. man whom I swore at/man at whom I swore
2''. man who swore at me
3''. man whose son swore at me

One argument that favors such an analysis comes from the fact that prepositions cannot be pied piped when that is the "relative pronoun." Since WH-pronoun preposing cannot prepose a lone preposition, it follows from the rule order WH-pronoun deletion—WH-pronoun preposing that such forms as #man at that I swore are impossible, whereas (1') and (1'') are permissible. No other analysis can handle this situation naturally. If that is taken to be a morphological variant of who/whom, then an ad-hoc condition must block the substitution of that just in case a preposition precedes the Wh-word.

23
If we imagine an analysis in which WH-preposing takes place first, followed by WH-pronoun deletion and that-introduction, we are faced with having to explain why the WH-pronoun fails to delete when preceded by a preposition.

It is also possible that from (i') the following string is derived by means of a rule of that-deletion.

i'''. man I swore at

This rule, it is often claimed, is the same that-deletion rule that occurs elsewhere in the grammar (e.g. The man believes (that) I swore at him). The matter is not so obvious, however. Consider the following two points: First, that-deletion, where the that is the head of a relative clause, is optional, and depends solely on whether a noun phrase follows the that directly in the string at the point where the rule applies. The underlying grammatical function of the "relative pronoun" that may even be subjective: The man that John believes is thinking of running for president →

The man John Ø believes is thinking of running for president.

The rule that deletes that-complementizers after embedding verbs must be a fuzzy-optional rule (Cf. Rodman (1972), where this fuzzy rule is discussed, and see Chapter Four for an explication of fuzzy rules). Its optionality is dependent on the embedding verb, as the data in (2.22) suggest.

(2.22) a. I believe (that) the man swore at me.
   b. I complained ?(that) the man swore at me.
   c. I regretted ?*(that) the man swore at me.
   d. I ask *(that) the man not swear at me.
Secondly, the that that introduces sentences in apposition to abstract head nouns is but fuzzy-deletable for many speakers, if it is deletable at all. Thus the string *The argument that John forgot* is ambiguous with the meanings suggested in (a) and (b):

(a) The argument which John forgot...

(b) The argument that John forgot is ridiculous—he has an excellent memory.

But the string *The argument John forgot* is unambiguous—interpreted with the (a)-meaning. The conclusion to be drawn, then, is that it is far from obvious that a single rule is involved in all cases of that-deletion.

To return to the main thread of the argument, consider the following ungrammatical sentences.

(2.23) a. *The reefer which the man who I charged five dollars for gave up his habit was a blend of dried rabbit pellets and oregano.

b. *The reefer that the man who I charged five dollars for gave up his habit was a blend of dried rabbit pellets and oregano.

There is no difference between (2.23a) and (2.23b)—both are equally ungrammatical. Yet only the a-sentence is blocked by the CNPC. Here is how the rules relevant to relative clause formation apply (on the higher cycle):

**WH-pronominalization:** The reefer that the man who I charged five dollars for WH-it gave up his habit was a blend of...

**WH-pronoun deletion** Does not apply.
WH-preposing: BLOCKED because WH-it cannot leave the complex NP \[NP[NP[the\ man]_{NP}g[who\ I\ charged\ five\ dollars\ for\ WH-it]_{NP}\]

But consider the b-sentence in the same way:

WH-pronominalization: Same as above

WH-pronoun deletion: The reefer that the man who I charged five dollars for \(\emptyset\) gave up his habit...

WH-preposing: Does not apply

The CNPC does not block this derivation.

These cases are in principle identical to the similar cases in Thai studied previously in this section. The conclusions that follow from these considerations of Thai and English can probably be replicated in any of the world's languages. I have done so for Mandarin Chinese and Korean (Rodman 1973), and Ross (1967) did so for English, Russian, Finnish and other languages.

Ross's investigations in this vein led him to reformulate the CNPC and the CSC in terms of the notion "island" or "syntactic island." An island is that part of a phrase marker that is dominated by a complex NP node as defined in the CNPC, or a coordinate node as defined in the CSC. I have illustrated a couple of islands in (2.24).
(2.24) John gave a dog and a cat to the girl who sells flowers.

Coupled with this notion of island is the following constraint.


The idea of a variable "crossing island boundaries" is used in an "intuitively clear sense" (p. 265), claims Ross. Since I will eventually define the notion "fuzzy island," I will be forced to give a rigorous formal definition of island, and I pass up the opportunity to do so at
this point. (Paradoxically, the more one deals with fuzzy matters, the more precise one is forced to be in his formulations -- perhaps because nothing fuzzy ever has an "intuitively clear sense.")

What Ross means by (2.25) is that there is no communication among elements of the phrase marker across island boundaries. A variable crosses an island boundary if it encompasses elements that lie on either side of the island boundary. The effect of (2.25) is that no rule is allowed to reorder any element in such a way that it crosses an island boundary, nor may any rule take effect if an island boundary intervenes between two non-variable elements of the structural description. Thus, for example, in (uni-directional) deletion under identity, if the controlling (i.e. ultimately non-deleted) element is on one side of an island boundary, and the element to be deleted is on the other side, the rule is blocked.

I'll illustrate the above by diagramming the ungrammatical Thai sentence (2.26), which we saw above as (2.19b).

\[ (2.26) \, ^*məːw \, ^\hat{t}ì: \, ðùːchaːj \, ^\hat{t}ì: \, ñēn \, ðçʰːb \, mː \]
\[ \text{cat} \quad \text{COMP} \quad \text{man} \quad \text{COMP} \quad \text{see} \quad \text{like} \quad \text{dog} \]
\[ \text{tua} \quad \text{nːn} \quad \text{dam} \]
\[ \text{classifier} \quad \text{that} \quad \text{black} \]
\[ ^*\text{The cat that the man who saw likes that dog is black.} \]
It is clear that this constraint blocks (2.23a) by virtue of its restriction of chopping rules, and (2.23b) by virtue of its restriction on either unidirectional rules of deletion or feature changing rules, whichever we consider the rule of WH-pronominalization to be.\textsuperscript{5}

2.2

2.2.1 The constraint given in (2.25) is totally within the spirit of the "Aspects" era of the Chomskian theoretical linguistics school in that it refers solely to the syntactic structure of phrase markers and the formal shape of grammatical rules. Thus given a transformational grammar of any language, it is an algorithmic matter to identify those strips
of phrase markers that qualify as islands and those rules whose action is constrained by (2.25). 6

Significantly, (2.25) is a constraint on rule application, and just as we found the CNPC and the CSC to be inadequate because they were confined to constraining rules of reordering, so we find (2.25) inadequate because it is confined to constraining grammatical transformations. I.e., we found constraints on derivations were necessary in cases where the CNPC and the CSC didn't hold, though the kind of constraint that seemed necessary was so strongly reminiscent of those two movement constraints that ultimately (2.25) had to be formulated, coalescing a large body of restrictions in the one constraint.

In an entirely analogous way, I shall note ungrammatical strings in Thai that appear as if their derivation should be blocked by (2.25), but because no relevant grammatical rule is involved in their generation, (2.25) cannot have the needed constraining effect.

Constituent Question formation in Thai is maximally simple: one merely puts the appropriate question word in the constituent-slot to be questioned. There is no reordering of elements, no feature changing and no deletion. The sentences (2.27) provide an illustration.

(2.27) a. khâw kin ?araj
    he eat what

    What did he eat?
b. khun chʰa: waː phǒm hɛn khraːj
   you believe COMP I see who
   Whom do you believe I saw?

It is certain that Constituent Question Formation in
Thai is effected via the process of lexical insertion. If
we consider ourselves in an Aspects framework, then such
questions are formed in the base. In particular, (2.25),
being a constraint on T-rules, has no jurisdiction over the
process. Nonetheless, in Thai it is not possible to question
elements of relative clauses, nor is it grammatical to ques-
tion elements in coordinate structures, as the data in
(2.28) and (2.29) show.

(2.28) a. phǒm chʰːb dɛːŋ thiː: phuːchaːj hɛn
       I like boy COMP man see
       I like the boy that the man saw.

b. "khun chʰːb dɛːŋ thiː: khraːj hɛn
   you like boy COMP who see
   "Whom do you like the boy that saw?

(2.29) a. khʰːw chʰːb huːmahːm lɛ phuːkhun chʰːb tæːnmːoː:
       he like onion and you like watermelon
       He likes onions and you like watermelon.

b. "khʰːw chʰːb huːmahːm lɛ phuːkhun chʰːb ʔaraj
   he like onion and you like what
   "What does he like onions and you like?

c. "khʰːw ʔaraj lɛ phuːkhun chʰːb tæːnmːoː:
   he like what and you like watermelon
   "What does he like and you like watermelon?

Some kind of statement is needed in Thai to block
sentences like these, but it is immediately obvious that any
such statement would have to be so similar in nature to (2.25) as to cast strong doubts on its existence as an independent statement. The existence (or non-existence, as it were) of (2.28) and (2.29) comprise damaging evidence against (2.25).

2.2.2 Our second argument will consider cleft sentences in English and Thai. Consider first Thai, and the sentences (2.14), which I reproduce here as (2.30).

(2.30) a. khâv khâ:j phâ:
he sell cloth

He sells cloth.

b. kʰː phâ: nâmːe thî: khâv khâ:j
particle cloth FOCUS COMP he sell

It's cloth that he sells.

(2.30b) is not derived directly from (2.30a) -- there is an intervening step that looks very much like a pseudocleft. 7

Roughly, then, we have:

```
          S
         /\        \   \  \
        NP   VP    khy:\  
           /\           /\        \  
          S  NP     be     thî:
             /\         \       /\     
            khâv  V    khâ:j  NP
               /\     /\      /\    /\  
              phâ: nâmːe  FOCUS  nâmːe
```

\[\text{RULE OF PSEUDOCLEFT}\]
(2.31) thî: khâw khâ:j khy: phâ:
COMP he sell be cloth

What he sells is cloth.

What happens at this point my investigations have not revealed in detail. Either the sentential subject is extrapolosed to the right, or the emphasized NP (phâ: "cloth" in this case) is reordered to the left. My guess is for extrapolosition. The particle kî: that shows up in the surface structure of the cleft functions as a linking element and is sometimes translatable as "then" or "consequently," but often is untranslatable. It does not correspond to the it that extrapolosed sentences in English leave behind. Its source, in any event, is unclear to me. The focus element nânlè, not present in the surface structure of pseudoclefts, does show up in the cleft. Its position to the right of the noun is normal for Thai: noun modifiers of all types follow their head noun:
(2.32) a. mae:w dam hâ: tua
cat black five classifier
five black cats
b. dég khon ní:
boy classifier this
this boy

khý: "be" is apparently deleted. This is not unusual for
Thai. The various copulas in the language are ephemeral,
excelling, it would seem, in optional appearance. Thus:

(2.33) ní: khy: sâmúd = ní: sâmúd
this be notebook

Here is the notebook.

bâ:n ní: khy: bâ:n phóm = bâ:n ní: bâ:n
house this be house I phóm

This house is mine.

The rule of Cleft Extraposition, then, operates as shown in
(2.34) to give the result shown in (2.35).

(2.34)

```
   S
  /  \
NF   VP
  |   |
S   V
  |   |
THI: NP   V
  |   |
kâ:w kâ:j
  |   |
KHY: N   NP
  |   |
PHA: FOCUS
```

34

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
RULE OF CLEFT EXTRAPOSITION

\[ S \]

\[ k^\circ: \quad NP \quad \text{FOCUS} \quad \text{thi}: \quad NP \quad VP \]

\[ \text{pha}: \quad n\text{n\text{\rlap{l}}}l\text{\rlap{\l}}}e \quad \text{kh\text{\rlap{\w}}w} \quad \text{kh\text{\rlap{\w}}}j \]

(2.35) \( k^\circ: \quad \text{pha}: \quad n\text{n\text{\rlap{l}}}l\text{\rlap{\l}}}e \quad \text{thi}: \quad \text{kh\text{\rlap{\w}}}w \quad \text{kh\text{\rlap{\w}}}j \)

particle cloth FOCUS COMP he sell

It's cloth that he sells.

Significantly, there are no complex noun phrases in the clefted structure, therefore no islands as far as (2.25) is concerned. Nonetheless, the embedded S, which functions as the "out-of-focus" clause, behaves like an island. Elements cannot be reordered out of it, nor can elements be questioned that belong to it.

(2.36) \( k^\circ: \quad \text{pha}: \quad n\text{n\text{\rlap{l}}}l\text{\rlap{\l}}}e \quad \text{thi}: \quad \text{khra}j \quad \text{kh\text{\rlap{\w}}}j \)

particle cloth FOCUS COMP who sells

*Who is it cloth that sells.

Again the restriction is so reminiscent of the island constraint that one feels compelled to include it in some broader based constraint.

This result can be duplicated in English. Pseudoclefts are formed from underlying structures such as (2.37) to give structures such as (2.38).
(2.37)

```
(2.38)

And the rule of Cleft Extraposition (Akmajian (1969)) applies to give (2.39).  

36

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(2.39)

It is possible, as Emonds (1969) suggests, that the extraposed S-node is embedded under the VP. This is irrelevant to our purposes, for the point here, as it was in our consideration of Thai, is that there are no islands. Nevertheless, the extraposed S behaves like an island, so, for example, the NP woman cannot be relativized:

(2.40) #The woman that it was food that Max gave to was very grateful.

One possibility here is that the rules of Pseudocleft and Cleft Extraposition are post-cyclic, and follow WH-REL in the ordering. Then Relative Clause Formation would not be allowed to apply to the tree structure (2.37) (perhaps a slight extension of Ross's Island Constraint). One might also argue that if it happens that Pseudocleft precedes WH-REL in the ordering, but Cleft extraposition follows it, then the Sentential Subject Constraint (Ross (1967):134) would prevent the NP woman from being relativized.

None of these arguments suffice to explain the ungrammaticality of (2.40), because it can be shown that Cleft
Extrapolation (hence Pseudocleft) is cyclic, and since the formation of relative clauses takes place on the following cycle, the rule ordering is:

Pseudocleft
Cleft Extrapolation
WH-REL

The argument is simple. Cleft Extrapolation must precede the known cyclic rule of Subject to Object Raising, i.e. the rule that derives (2.41b) from (2.41a):

(2.41) a. I believe I was stoned at Bob's party.
   b. I believe myself to have been stoned at Bob's party.

But now consider (2.42).

(2.42) I believe it to have been food that Max gave to the woman.

(2.42) shows that Cleft Extrapolation must precede Subject to Object Raising, and is hence cyclic, and must hence precede WH-REL, which takes place (at the earliest) on the following cycle. Pseudocleft, of course, must precede Cleft Extrapolation. The abbreviated structure in (2.43) sums up the argument.

(2.43)
We have yet again a restriction that is not covered by Ross's Island Constraint (2.25), but looks as though it ought to be.

2.2.3 A third argument against (2.25) is that it is incapable of accounting for clines of grammaticality (i.e., fuzzy data). Consider the following set of five facts:

(I) Subjects always have stronger membership (to be made precise in chapter three) in islands than non-subjects:
(2.44a) is less grammatical than (2.44b); (2.44c) is less grammatical than (2.44d) (henceforth (2.44a) < (2.44b), etc.).

(2.44) a. Who do you believe that kicked John?
   b. Who do you believe that John kicked?
   c. The man whom it was food that gave to the woman was named Max.
   d. The woman whom it was food that Max gave to was very grateful.

Similarly, in Mandarin Chinese it is not entirely ungrammatical to question non-subject elements of relative clauses, but it is measurably worse to question subjects. (Mandarin Constituent Question Formation is identical (in its grosser aspects) to Constituent Question formation in Thai. I.e., a question word (shème for non-human, shéi for human) is placed in the NP slot to be questioned--there is no reordering of elements, no deletion and no feature-changing.) Thus (2.45a) > (2.45b).
(2.45)  a.  tā xǐhuān dài shémma de háizi
       he like wear what rel. part. child
       What does he like the child who is wearing?

       Cf.

       tā xǐhuān dài yǎnjìng de háizi
       he like wear glasses rel. child

       He likes the child who is wearing glasses.

       b.  tā xǐhuān shéi dài de yǎnjìng
           he like who wear rel. glasses
           Who does he like the glasses that is wearing?

       Cf.

       tā xǐhuān háizi dài de yǎnjìng
       he like child wear rel. glasses

       He likes the glasses that the child is wearing.

(II) Contrary to results predicted by (2.25), it seems that many speakers allow the islands formed by sentences in apposition to abstract head nouns to be penetrated. Thus (2.46) (also (1.2)) is semi-grammatical for many speakers.

(2.46) It was Sally who he invented a story that John ran over.

One of the several interesting things about such cases is that there appears to be a grammaticality gradient that is related to the determiner on the head noun of the N-S construction. Thus (2.47a) > (2.47b) > (2.47c) > (2.47d) > (2.47e).

(2.47)  a.  It was Sally who he invented a story that John ran over.

       b.  It was Sally who he invented a certain story that John ran over.
c. It was Sally who he invented the story that John ran over.

d. It was Sally who he invented this story that John ran over.

e. It was Sally who he invented his own story that John ran over.

This determiner hierarchy of "degree of identifiability" (see section 3.1.3.1.1.2) will show up elsewhere in the grammar, as we shall see in chapter three. It affects the strength of islands.

(III) In several of the languages that I have investigated, it appears that rightmost NP's in an NP coordinate structure have the weakest island membership. (This fact, and the facts stated in (I) are probably not independent—certainly not in SVO languages.) For example, in Thai, it is better to question the rightmost "watermelon" in (2.48) than to question "onions," i.e., (2.48b) > (2.48c).

(2.48) a. kháw chò:b huáh:y:m láe taengmo:
    he like onion and watermelon
    He likes onions and watermelon.

b. kháw chò:b huáh:y:m láe ?àraj
    What does he like onions and?

   c. kháw chò:b ?àraj láe taengmo:
    What does he like and watermelon?

   These results, significantly enough, are reproducible in Korean. Like Thai and Mandarin, Korean effects the formation of constituent questions by the placing of an appropriate question word in the appropriate, unregulated slot.

In (2.49), b > c.

41
(2.49) a. no-nin sakwa wa kwaca-lal cohaha-nin-ta
   you apples and cookies like
   You like apples and cookies.

   b. no-nin muos-wa kwaca-lal cohaha-nin-nya
      you what-and cookies like
      What do you like and cookies?

   c. no-nin sakwa - wa muos-lal cohaha-nin-nya
      you apples and what like
      What do you like apples and?

Yet further evidence in this vein comes from the other side of the world, strengthening further my belief that whatever is going on here is universal. Russ Schuh has advised me that data entirely analogous to the Thai and Korean data just presented can be found in Ngizim, a Chadic language. Constituent Question Formation in Ngizim is effected just as in Thai, Mandarin, etc., i.e., probably not by grammatical rule. In Ngizim, however, it is ungrammatical to question elements that belong to relative clauses, and it is ungrammatical to question any but the rightmost NP of a conjoined series of NP's.

(2.50) a. #kà nci tam naa albasañ
   you like what and onions
   What do you like and onions?
   kà nci albasañ naa tam
   What do you like onions and?

And even in English there is a dialect$^9$ in which it is possible to question the rightmost NP of a series of conjoined NP's.
(2.51) a. Who did you see John and?
   *Who did you see and John?
b. What did you eat peaches and?
   *What did you eat and peaches?

(IV) In English different types of complements behave differently with respect to one another insofar as they permit their elements to be reordered and otherwise affected by grammatical rules. We might say that the complements behave as if they were very weak islands, with some stronger than others. One of the weakest islands is a that-S complement with the that deleted. It even allows subject NP's to be affected by external influences. As the complement type changes, island strength may vary somewhat. In (2.52), corresponding sentences in (2.52a) ≥ (2.52b) ≥ (2.52c) ≥ (2.52d) ≥ (2.52e) ≥ (2.52f).

(2.52) a. What do you think John smokes?
   Who do you think smokes pot?
   Cf.: I think John smokes pot.
b. What do you think that John smokes?
   Who do you think that smokes pot?
   Cf.: I think that John smokes pot.
c. What did you plead for John to smoke?
   Who did you plead for to smoke pot?
   Cf.: I pleaded for John to smoke pot.
d. What do you abhor John's smoking?
   Left Branch Condition affects this datum.
   Cf.: I abhor John's smoking pot.
e. What did you criticize John's smoking of?
Left Branch Condition affects this datum.
Cf.: I criticized John's smoking of pot.

f. What did you publicize John's purchase of?
Left Branch Condition affects this datum.
Cf.: I publicized John's purchase of pot.

It seems to me that these judgments of grammaticality are basically correct, and that any island constraint that purports to account adequately for the facts of English will have to deal with such data.

(V) For some years it has been known that factive clauses are somewhat island-like in their behavior.

(2.53) a. John thought that Mary bumped into Godzilla's roach clip.
What did John think that Mary bumped into?

b. John realized that Mary bumped into Godzilla's roach clip.

What did John realize that Mary bumped into?

Not only do the complements to factive predicates resist chopping rules, but subject to object raising cannot take place out of a factive clause, nor can the element neg be transported out of a factive clause (all this discussed in section 2.3.3.3).

The solution to this problem in terms of Ross's Island Constraint is to posit an underlying head noun the fact for factive clauses, and order its deletion late enough so that island sensitive rules have all been passed by the time it
occurs. The fact that S presents the facade of a complex NP and accounts for the island-like properties that are observed.

Besides the spurious the fact analysis, which will be dealt with in section 2.3.3.3, there are two other important problems. One is that sentences like What did John realize that Mary bumped into are simply not as ungrammatical as sentences like What did John realize the fact that Mary bumped into. Actually, one can choose factives whose complements behave like extremely weak islands (cf. (2.54d) below)—but S's in apposition to the fact always behave like fairly strong islands.

In the second place there are clines of island-strength among the complements of various factives: (2.54d) > (2.54e) > (2.54f).

(2.54) a. I am sorry that the baby crapped in your hand.
   b. I regret that the baby crapped in your hand.
   c. I deplore that the baby crapped in your hand.
   d. Whose hand are you sorry that the baby crapped in?
   e. Whose hand do you regret that the baby crapped in?
   f. Whose hand do you deplore that the baby crapped in?

Here again we have a situation that cannot in principle be handled by a constraint in a discrete (i.e. non-fuzzy) grammar.
2.3

2.3.1 In an attempt to provide a somewhat explanatory account for some of the phenomena associated with island constraints, we might look at one of the stumbling blocks of the CNPC. We noted above in section 2.2.2 that the out-of-focus clause in a cleft construction exhibits island-like properties, both in English and in Thai. In particular, you cannot focus an element in the out-of-focus clause by means of a second clefting operation. The sentences in (2.55) are ungrammatical.

(2.55) a. *The one who what John told he would try to eat were beans was Mary.

b. *The one who it was beans that John told he would try to eat was Mary.

c. *It was Mary who it was beans that John told he would try to eat.

Cf.:

(2.56) a. John told Mary he would try to eat beans.

b. What John told Mary he would try to eat were beans.

c. It was beans that John told Mary he would try to eat.

d. The one who John told he would try to eat beans was Mary.

e. It was Mary who John told he would try to eat beans.

It is necessary to show that the ungrammaticality of sentences (2.55) does not follow from other considerations. For instance, any string in (2.55) could have the underlying structure shown in (2.57).
It has already been shown that the clefting rules are necessarily cyclic, so the output at the end of the $S_1$-cycle might be either (2.58a) or (2.58b).
(2.58) a.

In either case (2.25) will not prevent reapplication of the clefting rules on the \( S_0 \) cycle to give the ungrammatical (2.55).

It might, on the other hand, be argued that there is a constraint at work in English something to the effect that it
is never allowed to displace more than one constituent from its slot in deep structure. This transparently perceptual constraint has a lot going for it, actually, and will play a role in several of our conclusions in this thesis. There are difficulties, however. If one considers passives to be derived by rule, then the fact that one can question, cleft, topicalize, etc. constituents in passive constructions would be damaging counter-evidence. Also, a sentence like Who are letters tough for John to write to, which I admit is not fully grammatical, is nonetheless not nearly as bad as (2.55), and also raises difficulties with this potential constraint. And insofar as multiple clefting is concerned, the near perfect grammaticality of (2.59), with its two displaced constituents, forces us to seek other means of accounting for the ungrammatical (2.55).

(2.59) It was Mary who John told it's beans he would try to eat.

The one who John told what he would try to eat were beans was Mary.

We may be led to those other means by considering clefting from a somewhat more semantic point of view. First, consider a clefted structure itself. It divides a sentence into two parts, one of which, the focus, is ascribed augmented communicative importance, and the other of which, the out-of-focus or presupposed clause, is ascribed diminished communicative importance. Following Schachter (1973), we might call these parts, respectively, the foreground and the background.
Now consider clefting from an operational standpoint. It augments an element in communicative importance, that is, it foregrounds the element, while the rest of the clause is concomitantly backgrounded. Clefting, then, is a foregrounding operation.

A plausible hypothesis follows from these considerations, and was made by Schachter (1973:44f):

I would like to suggest that the relevant constraint may have to do with what I have called foregrounding. Thus it may be the case that, once a sentence has been transformed into a foreground-background sequence, this sequence is immune from further transformational division into foreground and background.

It is clear that (2.55c) is ungrammatical in terms of this constraint because the sentence It was beans that John told Mary he would try to eat is not subject to another foregrounding operation that would place the once backgrounded Mary into the foreground.

Apart from accounting for a goodly number of syntactic phenomena, especially after we have modified it slightly, which will be done forthwith, this constraint has the added advantage of being explanatory in terms of linguistic cognition. As Schachter notes: (1973:45):

Such a restriction might be described in terms of what Bever 1970 [The Cognitive Basis for Linguistic Structures. Cognition and Language Learning, ed. by J. R. Hayes. New York:Wiley] has called a perceptual strategy. It could be proposed, for instance, that once a constituent has been identified as the foreground of one construction, it cannot, on the basis of evidence appearing at a later point in the sentence, be identified as (part of) either the foreground or the background of another construction.
Here then is a constraint that avoids the criticism levelled against (2.25) that it lacks explanatory value. Moreover, the constraint appears to have the power to account for some of what (2.25) couldn't account for. We will, then, undertake a careful investigation of this constraint to see if it is indeed what we seek.

Consider first the rule of Object Preposing, often called Topicalization, or more ethnically Yiddish Movement. It has been argued by Lakoff (1969) and Postal (1971) that Object Preposing in English is a post-cyclic rule that follows the WH-preposing rules. The arguments are cogent and we will assume their validity. Object Preposing is undisputably a foregrounding operation, and as such it is subject to the constraint on foregrounding suggested by Schachter. We therefore predict the ungrammaticality of (2.60b).

(2.60) a. It's Mary that John told he likes beans.

b. *Beans, it's Mary that John told he likes.

An analogous prediction is true in Thai:

(2.61) a. kô: sã:khā: nãnlâe thî: phôm bɔː:g
particle Sakaā FOCUS COMP I tell
wâ: thə: sỹ: nãngsỹ: lêm nĩ:
COMP she buy book classifier this

It's Sakaā that I told she bought this book.

b. *nãngsỹ: lêm nĩ: sã:khâ: nãnlâe thî:
book class. this Sakaā FOCUS COMP
phôm bɔː:g wâ: thə: sỹ:

*This book, its Sakaā that I told she bought.
We also noticed in section 2.2.2, as part of our criticism of Ross's Island Constraint (2.25), that in English elements in the out-of-focus clause of a cleft construction are not subject to relativization. We might note in passing that this is also true in Thai—a fact as equally perplexing for (2.25) as the English example (2.40).

(2.62) a. kʰ: khruː nānlæ thiː kʰːj phː part. teacher FOCUS COMP sell cloth

It's the teacher who sells cloth.

b. *phː thiː (kʰː) khruː nānlæ thiː cloth COMP part. teacher FOCUS COMP

kʰːj phaeːng sell expensive

*The cloth that it's the teacher who sells is expensive.

Again the foregrounding constraint succeeds where (2.25) fails, for Relative Clause Formation is a foregrounding operation and may not operate on elements of the out-of-focus (backgrounded) clause. As argued in Schachter (1973), a relative construction divides a sentence into a head and an attribute, with the head receiving elevated communicative prominence (i.e., it's foregrounded), and the attributive clause being concomitantly ascribed lesser communicative importance (i.e. it's backgrounded). That the entire relative construction functions syntactically as a nominal further supports the contention that the head noun is the recipient of the high communicative importance expected of a foregrounded element.
One cannot relativize out of a cleft construction for precisely the same reason that one cannot topicalize or re-cleft; elements once backgrounded cannot be foregrounded.

Turning things around, now, we find ourselves with a constraint that also predicts that elements in relative clauses cannot be clefted, topicalized or relativized, for that violates the foregrounding constraint too, since a relative structure is a foreground/background structure. In this case the constraint has impinged on territory that one normally thinks of as belonging to (2.25). Having made the breech, why not push on to test the viability of replacing (2.25) with the foregrounding constraint?¹⁰

A test case that immediately comes to mind has to do with the formation of constituent questions in Thai. It was partly on the basis of the interaction between the processes of Object Preposing, Clefting and Relativization with Constituent Question Formation that led us to doubt (2.25) to begin with (see section 2.2.1). I would like to suggest that constituent questions put into the foreground of the sentence one (or more) constituent slots, and impose the meaning "?" on that slot. The remainder of the sentence is backgrounded. The foregrounding constraint predicts the ungrammaticality of strings like (2.28b) (questioning inside a relative clause) and (2.36) (questioning inside the out-of-focus clause of a cleft sentence) as well as strings like (2.63), where it is shown that in Thai it is ungrammatical

53
to question elements in a structure that is to undergo Object Preposing.

(2.63) *nom kâːw niː khāw chya waː khraj
dyːm (Cf. 2.7b)
milk glasses of this he believe COMP who drink

The island constraint (2.25) fails to predict ungrammaticality in any of these cases.

We have observed that the rules of Relative Clause Formation, Constituent Question Formation (be it a rule or not), Clefting and Object Preposing are all foregrounding operations, and that relative constructions, constituent questions, cleft constructions and topicalized sentences are all strings that have been divided into a foreground and a background. The constraint on foregrounding suggested in Schachter (1973) restricts the imposition of one of these structure types on another in a way that is in accord with the observed data.

The foregrounding processes (where "process" does not always imply a grammatical rule, e.g. constituent question formation in Thai), besides being constrained from applying to structures that have already undergone a foreground/background division, are also restricted in their application to coordinate structures in a way that we are familiar with from having discussed the CSC.

Let us view coordinate structures as the division of semantic material in a sentence into parts of equal communicative importance. It appears that there is a constraint to
the effect that once a division is made into equal parts, no
element belonging to any one part (conjunct), nor any part
(conjunct) itself, may be foregrounded at the expense of an
equal partner. This last clause is important because ele-
ments within a single conjunct may be foregrounded at the
expense of other elements in the same conjunct, as the data
in (2.13) showed.

We may broaden Schachter's constraint, now, into a
more encompassing statement. We use the term "subdivision"
to mean a foreground/background dichotomy, or a division in-
to equal parts by means of coordination.

(2.64) A sentence, or part of a sentence, once
subject to subdivision, may not have further
subdivisional relationships imposed on it.

We hypothesize that (2.64) replaces (2.25) (but see
footnote 10), as well as Schachter's principle on which it
is based.

2.3.2

2.3.2.0 So far we have discussed in some detail the follow-
ing situations that can be taken as favoring (2.64) over
(2.25):

1) Constituents that belong to relative clauses may not
be questioned in Thai.

2) Clefts in English and Thai behave as if they contained
islands, though they do not as far as (2.25) is
concerned.

3) Object Preposing in English or Thai leads to un-
grammatical results if a constituent has been questioned.
In the following three sections we shall look at further evidence that favors (2.64) over (2.25).

2.3.2.1 Pretending to forget the "rightmost" exceptions noted above in section 2.2.3 (III), we see that (2.64) predicts the ungrammaticality of questioning elements in a coordinate structure in Thai, which (2.25) cannot do.

(2.65) a. *khun chā b tāraj lāe tae ngmo: you like what and watermelon

*What do you like and watermelon?

b. khraj d’y:m nom kā:w nì: lāe who drink milk glasses of this and

nāgrīan d’y:m cha: thūaj nì: student drink tea cup of this

*Who drank this glass of milk and the student drank this cup of tea?

c. khru: d’y:m nom kā:w nì: lāe teacher drink milk glass of this and

khraj d’y:m cha: thūaj nì: who drink tea cup of this

*The teacher drank this glass of milk and who drank this cup of tea?

2.3.2.2 While (2.25) prevents elements from being chopped out of a relative clause that is intact, if the clausal part of the relative clause is extraposed by the rule Extrapo-

sition from NP, thus rendering the "complex" NP no longer com-

plex (i.e. destroying the island), it nonetheless remains true that the clause retains its island-like properties. The data in (2.66) illustrate this fact.
(2.66) a. A man who belongs to SDS was arrested on campus.

b. A man was arrested on campus who belongs to SDS. (Ex. from NP has applied)

c. *What organization was a man arrested on campus who belongs to?

d. *The organization that a man was arrested on campus who belongs to was SDS.

This appears at first blush to comprise support for (2.64) in that we can legitimately argue that the clausal part of a relative clause is backgrounded irrespective of its syntactic position in the sentence, and as background it is not subject to the foregrounding operations of relative clause formation or constituent question formation.

However this argument can be gotten around. Suppose, as Ross (1967) did, that ExNP is post-cyclic and follows the rules of WH-preposing in the ordering. Then the island will be intact at the critical moment and (2.25) will block the ungrammatical (2.66cd).

It will be recalled that the same kind of argument was countered in section 2.2.2 during the discussion of the clefting rules, where it was shown that those rules must be cyclic and an ordering argument was not able to explain the ungrammatical strings observed at that time. The matter is more difficult regarding ExNP. I know three arguments in favor of its cyclicity, none of which are particularly compelling. Let us briefly review them.
(i) In Ross (1966) it is argued that pronominalization is cyclic. ExNP must precede pronominalization as the following argument shows.

(2.67) a. The idea that Helen will be able to ride a bike pleases Helen.

b. The idea pleases Helen that Helen will be able to ride a bike.

The sentences (2.67ab) are related by the rule of ExNP.

Suppose this rule did not precede Pronominalization. Then from (2.67a), (2.68a) can be derived by backwards pronominalization, the proper "command" relationships being met, and (2.68b) can be derived by ordinary left to right pronominalization.

(2.68) a. The idea that she will be able to ride a bike pleases Helen.

b. The idea that Helen will be able to ride a bike pleases her.

Now let ExNP apply:

(2.69) a. The idea pleases Helen that she will be able to ride a bike.

b. *The idea pleases her that Helen will be able to ride a bike.

Very ad-hoc constraints would be needed to block the application of ExNP to (2.69b)—in fact, the conditions on backwards pronominalization would need to be incorporated into that very non-pronominalization-like rule.

On the other hand, if the rule ordering is ExNP Pronominalization, (2.69b) is automatically blocked because the condition on backwards pronominalization is not met by
(2.67b). Since ExNP precedes the cyclic rule of pronominalization, it must be cyclic.

Lakoff and others have recently argued, however, that pronominalization is post-cyclic, and extremely late in the ordering, so to the extent that their arguments are valid, the argument in (i) is vitiates.

(ii) Emonds, in his dissertation (1970), shows that ExNP is a structure preserving transformation, and since only root transformations are post-cyclic in Emonds system, to the extent that the Structure Preserving hypothesis is valid, there is an argument in favor of the cyclicity of ExNP.

(iii) Chomsky (1971) submits the following hypothesis as a general condition on linguistic theory, which is needed "to further sharpen the notion 'transformational cycle'."

(2.70) No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

Observe the application of ExNP to (2.71) giving (2.72), modelled after what I think would be Chomsky's analysis stripped to the bare essentials for purposes of expostulation.

(2.71) That a man who belongs to SDS was arrested on campus frightened Kathy.
That a man was arrested on campus who belongs to SDS frightened Kathy.

Note that the operation is confined to the next to last S-cycle. In fact, (2.73) is not possible.
(2.73) *That a man was arrested on campus frightened Kathy who belongs to SDS.

If ExNP is post-cyclic, it will, through the application just illustrated, violate condition (2.70). Hence ExNP is a cyclic rule.

If ExNP is post-cyclic, then both (2.25) and (2.64) will suffice to block ungrammatical strings such as (2.66cd). But if ExNP is cyclic, because relative clause formation in a structure that might underlie (2.66d) occurs on a later cycle, (2.64) will work where Ross's (2.25) will fail.

2.3.2.3 I noted at the end of section 2.3.1 that in a co-ordinate structure, division into parts of equal communicative importance has taken place, and it is impossible to foreground any element of one conjunct, or the conjunct itself, at the expense of backgrounding elements in an equal partner. In English, this restriction, included in the statement (2.64), and the CSC often cover the same ground. Ross (1967: 96f) noted a peculiar exception to the CSC. An element can be chopped out of a coordinate structure if it occurs in each conjunct of the coordinate structure, and if it is operated on by an "across-the-board" rule. Relative Clause Formation is such a rule, and its ability to operate "across-the-board" is exemplified in (2.74) (Ross's 4.122), p. 98).

(2.74) Students who fail the final exam or who do not do the reading will be executed.

It is significant that the principle (2.64) does not militate against across-the-board relativization. Since the
same element is foregrounded out of each conjunct, the status
of no conjunct is altered relative to an equal partner; the
conjuncts all become background material together, and re-
tain their equality. If a single conjunct fails to join
its partners in background status, (2.64) will block the
derivation, as it should. (But see 3.2.2.2:2 for a reanaly-
sis of the across-the-board phenomenon.)

(2.75) *Students who fail the final exam or who do
not do the reading or professors avoid giving
coherent lectures will be executed.

Thus (2.64) explains a puzzling counter-example to the CSC,
which provides further motivation for adopting it in the
place of (2.25).

2.3.3

2.3.3.0 The following three sections discuss some facts
about English which I one time took to favor the existence
of (2.64) (Rodman (1973)). In each of the three cases I
observed island-like behavior in certain types of structures.
I then argued that foregrounding and/or backgrounding played
a role in the formation of the structure, and hence accounted
for why the structures behaved like islands. In each of
these three cases it was not clear to me that the foreground/
background dichotomy existed, and I began to wonder whether
other factors were involved as well. There were, and they
will be discussed in chapter three.

2.3.3.1 Adverbial subordinate clauses (i.e. complements of
"subordinating conjunctions") behave as if subject to (2.25).
(2.76) a. The princess threw herself at the huntsman's feet \{ because after while although \} he kissed her hand.

b. *What part of her body did the princess throw herself at the huntsman's feet \{ because after while although \} he kissed?

If we naively take the word "conjunction" at face value, we might try to account for (2.76b) by the CSC (as embodied in (2.25)) or by (2.64). This plainly won't work, however, because the non-subordinate main clause does not behave like an island.

(2.77) At whose feet did the princess throw herself \{ because after while although \} he kissed her hand?

Although it has been proposed to account for strings like (2.76b) by deriving adverbial subordinate clauses from underlying relative clauses (e.g., after he kissed her hand from after the time at which he kissed her hand or perhaps at a time that was after the time at which he kissed her hand), Emonds has argued convincingly (1970:137f) that adverbial subordinate clauses are simply complements to prepositions. He observes that only prepositions and verbs take direct objects, and by analyzing verbal particles as intransitive prepositions, he also observes that both prepositions and verbs may be intransitive. Since verbs take sentential complements it is reasonable to see if prepositions do so.
The answer is affirmative, for Emonds successfully defends the hypothesis that prepositions take sentential complements that appear in surface structure as adverbial subordinate clauses.

One of the arguments that favored the relative clause analysis of subordinating conjunctions (which, it should be noted, analyzed the subordinating conjunction itself as a preposition) was precisely its island-like behavior which (2.76b) exemplifies. By ordering the rule that deletes the superfluous structure after the WH-preposing rules, (2.25) can account for that behavior. If the analysis is streamlined in the way Emonds suggests, then something else will have to account for it. The proposal is to simply note that a subordinate clause, as its name suggests, comprises background material. Principle (2.64) prevents elements of background material from being foregrounded by rules of Constituent Question Formation, Relative Clause Formation and so on.

What I don't like about this solution is that it isn't as clear to me that subordinate clauses comprise background material as it is when, say, clefts are involved. Certainly one can imagine adverbial subordinate clauses that are in no way reduced in communicative importance—indeed, elliptical sentences often consist of a clause in isolation, or even a lone subordinating conjunction.

I conclude, then, that there are no facts about subordinating conjunctions that support (2.64). Any set of

64
principles that replace \((2.64)\) will hopefully account more
perspicuously for the observed data.

2.3.3.2 In part, the CNPC was designed to prevent elements
from being chopped out of sentences which are in apposition
to an abstract noun like \textit{fact}, \textit{claim}, etc. For example, the
strings \((2.78bc)\) are ungrammatical according to \((2.25)\).

\begin{align*}
(2.78) & \quad \text{a. The minister believed the claim that Alex}
\text{had been cured of his violent tendencies.} \\
& \quad \text{b. Who did the minister believe the claim}
\text{that had been cured of his violent tendencies.} \\
& \quad \text{c. The violent tendencies that the minister}
\text{believed the claim that Alex had been cured}
\text{of were lurking just beneath the surface of}
\text{his psyche.}
\end{align*}

In Rodman (1973) I claimed that constructions of this form
are made up of material divided into foreground and back-
ground. The abstract head noun serves as the foreground and
indicates the kind of sentence which is about to follow, be
it a fact, an assertion, a claim, a proposal, a denial, a
suggestion, etc. The sentence in apposition is reduced in
status to background material, and is therefore subject to
\((2.64)\), which is apparently as it should be.

The problems here are twofold. First, it isn't entire-
ly evident that the appositional sentence is truly back-
grounded material, though I believe it is to some extent.
The syntactic category of the whole \textit{N-S} construction is, af-
ter all, the same as the syntactic category of its head.
The problem, as we shall see, is that nothing else contri-
butes to the strength of the island in \textit{N-S} constructions,
which remains small. This implies, to look ahead, that the foreground/background dichotomy is just a single factor among several. (See chapter three, especially sections 3.1.3.1.2.)

Secondly, as I mentioned earlier in section 2.2.3, for many speakers the sentential complement to an abstract head noun does not behave like an island; some speakers accept (2.79) (given above as (2.46)).

(2.79) It was Sally who he invented a story that John ran over.

In fact, in Japanese, apparently most if not all speakers accept (2.80).

(2.80) kono booshi-o Max-wa (Mary-ga kabutte ita
this hat Max Mary wearing is
to iu uwasa-o shinjite iru
that rumor believing is

#This hat, Max believes the rumor that Mary wore.

The rule of Topicalization in Japanese is given by Akatsuka (1969:16) as (2.81).

(2.81) X - NP - Y
   1  2  3   ------→  OPTIONAL
   2#(1  Ø  3)

It relates sentences like (2.82a) to (2.82b).

(2.82) a. Alice-wa (Max-ga hon-o katta) to iu
       Alice Max book bought that say
       Alice says that Max bought a book.

       b. hon-o, Alice-wa (Max-ga katta) to iu
       A book, Alice says that Max bought.

Since "the speaker must place an unusually heavy stress on the preposed NP, and a pause between it and the following
constituent" (Akatsuka 1969:15), it seems safe to assume that Topicalization is a foregrounding operation and should be constrained by (2.64), as it is with respect to coordinate structures, but Akatsuka claims that sentences like (2.80) are nonetheless grammatical.

Two lines of investigation come to mind. First, do all foregrounding operations behave as Topicalization with respect to the N-S construction in Japanese, and second, are there any N-S constructions that show stronger island properties than the one exemplified in (2.80)? We shall investigate questions of this sort for English in section 3.3, but the matter is too delicate for a non-native, non-student of Japanese to take up with respect to that language.

Returning to English, we noted the existence of the cline of grammaticality given as (2.47), which both (2.25) and (2.64) are in principle incapable of handling. Here is another cline in which the matrix verb plays a role:

(2.83b) > (2.83d) > (2.83f) > (2.83h).

(2.83) a. The governor made a claim that he shouldn't have to pay taxes.

b. What did the governor make a claim that he shouldn't have to pay?

c. The governor believed a claim that he shouldn't have to pay taxes.

d. What did the governor believe a claim that he shouldn't have to pay?

e. The governor discussed a claim that he shouldn't have to pay taxes.
f. What did the governor discuss a claim that he shouldn't have to pay?

g. The people deplored a claim that the governor shouldn't have to pay taxes.

h. What did the people deplore a claim that the governor shouldn't have to pay.

2.3.3.3 Not unrelated to what we noted in the previous section are the island-like properties displayed by the sentential complements of factive predicates. Whatever the absolute grammaticality judgments, (2.84bc) are worse than (2.85bc).

(2.84) a. It is deplorable that Guinevere bit the baby.

b. The baby that it is deplorable that Guinevere bit now wets its diaper whenever it sees her.

c. It's Michael's baby that it is deplorable that Guinevere bit.

(2.85) a. It is unlikely that Guinevere bit the baby.

b. The baby that it is unlikely that Guinevere bit nonetheless wets its diaper whenever it sees her.

c. It's Michael's baby that it is unlikely that Guinevere bit.

In their pioneering work on factivity, the Kiparskys (1968) proposed that factive predicates be subcategorized to take factive complements. Factive complements have the head noun fact in remote structure. In part, they were motivated by the following paradigm.

(2.86) a. The fact that Guinevere likes the odor of hash is suspicious.

b. *The fact that Guinevere likes the odor of hash is likely.
Thus underlying (2.84a) is the string The fact that Guinevere bit the baby is deplorable, but underlying (2.85a) is That Guinevere bit the baby is unlikely. Grammatical rules are hypothesized that either delete the fact, giving That Guinevere bit the baby is deplorable, or convert the fact to it, giving It that Guinevere bit the baby is deplorable, with extraposition obligatory. Other rules may produce gerunds from factives.

One of the benefits that was thought to accrue from this analysis was that the CNPC, sensitive to the structure the fact that-S, would account for the island-like properties observed. The rule of Fact-deletion would merely need to be ordered after the rules of WH-preposing, Topicalization, etc. Before their manuscript went to press, however, the Kiparsky's added a footnote (9): "We thought earlier that the oddity of questioning and relativization in some factive clauses was also due to the Complex Noun Phrase Constraint: ...It appears now that questioning and relativization are rules which follow Fact-deletion."

The Kiparskys make this statement without qualifica-
tion, but it is, in fact, easy to demonstrate. In an earlier version of their 1973 work, Stockwell et. al. give the following paradigm in conjunction with an argument that the rule of Fact-deletion must precede Equi-NP-Deletion:

(2.87) a. I regretted the fact of my leaving/having left.
    b. *I regretted the fact of leaving/having left.
c. I regretted leaving/having left.
d. I regretted my leaving/having left.

(I have moved the asterisk from the (d) to the (b) sentence to correct an obvious typographical error.)

The authors reason that if the rule of Fact-deletion follows Equi-NP-Deletion, then Fact-deletion would be obligatory just in case the optional Equi-NP-Deletion had taken place. To avoid this complication, one merely orders the rule of Fact-deletion to precede the rule of Equi-NP-Deletion.

In his paper "Super Equi NP Deletion" from the Sixth Regional Meeting of the CLS, John Grinder gave reasons for ordering the rule of Equi-NP-Deletion before the WH-preposing rules (p. 306f). This, then, would imply that the rule of Fact-deletion precedes the WH-preposing rules.

Whether it's this argument, or some other that the Kiparskys had in mind, is not known to me. They raise several arguments in favor of their the fact that-S analysis, none of which, in my opinion, carry much weight. We might look briefly at some of them.

(i) Argument: Factive predicates all take complements of the form the fact that-S: non-factive predicates, in general, do not (cf. 2.86)).

Counterargument: The real generalization to be made here is that factive predicates take N-S complements indiscriminantly; any head noun, including fact, will do. Non-factives, on the other hand, must be subcategorized with respect to the head Noun of the N-S construction.
(2.88) a. I \{ believed \} \{ the fact \} that the dog barked.
Non-factive:
\{ asserted \} \{ the idea \}
\{ the proposal \}
\{ the claim \}
\{ the denial \}

b. I \{ resented \} \{ the fact \} that I had to write my thesis.
Factive:
\{ ignored \} \{ the notion \}
\{ the statement \}
\{ the suggestion \}
\{ the proposal \}

(ii) Argument: No factive predicate allows raising. This phenomenon is accounted for by the the fact that-S analysis in coordination with the CNPC.

Counterargument: Verbs must be marked in the lexicon as to whether or not they permit raising; non-factives like intimate, conjecture and conclude don't allow raising, while other non-factives like believe, suppose and allege do allow raising. Yet other non-factive predicates such as claim and maintain seem to have a fuzzy relationship with the rule of subject to object raising. There is, then a raising cline: (2.89) (2.90) (2.91).

(2.89) The council \{ believed \} the mayor's advisors to be crooks.
(2.90) The council \{ supposed \} the mayor's advisors to be crooks.
(2.91) The council \{ claimed \} the mayor's advisors to be crooks.

The generalization that factives never allow raising is expressable most naturally as the lexical redundancy: [+FACTIVE] --[ [-RAISING]. The superfluous syntactic structure is not needed.
(iii) Argument: No factive predicate allows neg-transportation. The element neg is prevented from transporting by the presence of the complex NP headed by the fact and the CNPC.

Counterargument: Same as in (ii). Verbs must be marked in the lexicon as to whether they permit neg-transportation or not, since some non-factives do, and some don't. (Or it's a lexical redundancy rule [+WEAK ASSERTIVE] → [+NEG-TRANSPORTATION] (see footnote 2). It's all the same to the present argument.)

(2.92) I \{believe\} he shouldn't go to the racetrack.
think
figure

= I don't \{believe\} he should go to the racetrack.
think
figure

It's \{likely\} he won't go to the racetrack.
possible

# It's not \{likely\} he'll go to the racetrack.
possible

I \{alleged\} he wouldn't go to the racetrack.
asserted
concluded

# I didn't \{assert\} he would go to the racetrack.
conclude

A further (counter) argument is embodied in the data of (2.93).

(2.93) a. I don't \{believe\} it'll budge an inch.
think
figure

\{he's got a plugged nickel.\}
b. *I didn't assert [it would budge an inch.] conclude he's got a plugged nickel.

In (2.93a) the neg must originate in the complement, since the idioms in the complement are grammatical only when negated. In (2.93b), the neg could not have been transported to the matrix verb; the complements were non-negative in underlying structure, hence the ungrammaticality.

Here, too, it is unnecessary to capture the generalization that factives predicates never allow neg-transportation. The lexical redundancy rule [+FACTIVE] → [-NEG-TRANSPORTATION] will suffice.

(iv) Argument: Pronominalization of the fact to it accounts for the presence of the non-expletive it in sentences like (2.94).

(2.94) I {deplore} {the fact} that bats suck blood.

Counterargument: In Stockwell, et. al. (1973), they state (p. 551): "But the Kiparskys' claim that it derives here from pronominalization of the fact is dubious in the extreme..." The authors present cogent and convincing evidence against deriving it from the fact in sentences like (2.94), and argue that it-that-S complements have a more direct source.

The Stockwell, et. al. work just cited is the only fairly comprehensive grammar I know that deals explicitly with the facts surrounding factivity. They find it necessary
to use lexical features such as \ [+FACT], [-FACT], [+/-FACT], 
for strictly factive predicates, strictly non-factive predi-
cates and predicates like report that function either as a 
factive or as a non-factive. These features carry all the 
semantic/syntactic information that could possibly be carried 
by the overt presence of the head noun the fact.

There are no reasons, then, to have an underlying fact 
in the complements of factive predicates. Therefore I con-
clude that the fact that-S analysis of the complements of 
factive predicates is spurious, and that the presence of the 
fact that-S has no special status. It is on par with the 
proposal that-S, the denial that-S, etc.

This leaves us with the problem of accounting for the 
reduced grammaticalness of (2.84bc). It was noted by the 
Kiparsky's that in a factive sentence such as (2.84a), the 
truth of the proposition expressed by the complement 
Guinevere bit the baby is presupposed by a sincere uttering 
of the entire sentence. The proposal I made in Rodman (1973) 
was that presupposed material is automatically backgrounded, 
allowing (2.64) to account for the lowered grammaticality of 
(2.84bc).

Such a hypothesis does not seem implausible to me.
Language, as a sophisticated system of communication, may be 
expected to give new and unexpected information precedence 
over old and known information. One way of doing that is to 
demote presupposed material in communicative importance. It
is, of course, true that new information is carried by presupposition in certain cases. For instance, a speaker may choose to communicate the truth of the proposition "Guinevere bit the baby" by uttering (2.84a) to a person who he knows isn't aware of Guinevere's penchant for tasting homo-sapien new-borns. But both participants in such a conversation will be aware (consciously or subconsciously) that that particular morsel of information has been communicated indirectly, or by implication—such a feeling arising from the fact that new information has been communicated in an "old-information" way. 13

Moreover, if we examine other instances where we have good reason to believe backgrounding has occurred, we almost always find presupposition.

It has been known for a long time that the truth of the proposition expressed by the out-of-focus clause of a cleft construction is presupposed, just as it is backgrounded. In a constituent question like Who ate beans?, it is presupposed that "someone ate beans," in accord with the claim I made earlier that the unquestioned part of the utterance comprises background material.

One is tempted, then, to test the hypothesis that presupposition material and background material are always co-incident. This turns out to be untrue.

There is background that is not presupposed. In the utterance Sopwith wants to find the camel that trampled his
sweetpeas, the backgrounded *a camel trampled Sopwith's sweetpeas* corresponds to the presupposition "some camel trampled Sopwith's sweetpeas." But in a slightly different mode of relative clause such as *Edgar wants to find an elephant that flies*, the backgrounded *an elephant flies* has no such presuppositional correspondence since an uttering of the sentence does not involve the presupposition "some elephant flies."¹⁴

The ability of (2.6⁴) to account for the lowered grammaticality of (2.8⁴bc) hinges on the hypothesis that presupposed clauses are also backgrounded--a plausible, though hardly well-established idea. What ultimately defeats (2.6⁴) is that it cannot possibly account for the grammaticality cline in (2.5⁴) or for the relatively slight degree of lowered grammaticality that one perceives in (2.8⁴bc)--much slighter than what one perceives in (2.9⁵).

(2.9⁵) The baby that Michael deplored the fact that Guinevere bit now wets its diaper whenever it sees her. (Cf. (2.8⁴b))

In chapter three we will correct this fundamental weakness. (See especially section 3.1.3.1.3).

2.3.4

2.3.4.1 In this and the following two sections I shall discuss factors that militate against (2.6⁴) as a principle of grammar. Indirect question complements (or Q-complements) behave as if they were islands. They do not fall under the control of (2.2⁵), and Ross did not discuss them in his
thesis. The ungrammatical strings in (2.96) are far too 
reminiscent of island constraint violations (or violations 
of (2.64)) to warrant a separate constraint.

\[(2.96) \quad \begin{align*} 
\text{a.} & \quad \begin{cases} 
\text{know} & \text{why Robin doesn't keep the money.} \\
\text{asked John} & \text{how many flowers Robin sent Lady Marian.} \\
\text{wonder} & 
\end{cases} \\
\text{b.} & \quad \begin{cases} 
\text{know} & \text{why Robin doesn't keep?} \\
\text{asked John} & \\
\text{wonder} & 
\end{cases} \\
\text{*Who did you ask John how many flowers Robin sent?} \\
\text{*Who do you (know wonder) how many flowers Robin sent?} \\
\text{c.} & \quad \text{It is a (matter question) of who will bell the cat.} \\
\text{d.} & \quad \text{*What is it a (matter question) of who will bell.}
\end{align*} \]

If principle (2.64) could explain these data, such an 
explanation would go a long way in verifying its existence.
A cursory look at a sentence such as Holmes wondered who 
would eat the next victim might lead us to suggest that with-
in the complement of wonder there is a foreground/background 
dichotomy, where who is the foregrounded element, just as 
in ordinary WH-questions. Someone would eat the victim 
would be backgrounded (and presupposed), and it would not be 
possible, say, to relativize the NP position occupied by vic-
tim and get The victim that Holmes wondered who would eat...
Such a suggestion is supported by the fact that within the 
complement, the usual constraints obtain:
(2.97) a. *I wonder who John saw the car that ran over.
   b. *I wonder where John went to Paris and.
   c. *I don't know why John told Mary that the
      bartender he had talked to the other night
      had a scarred cheek.\textsuperscript{15}

I believe, then, the explanation provided by (2.64) is
valid \textit{in part} for all types of embedded questions but one:
(See section 3.1.3.1.3 for further discussion.)

(2.98) a. I wonder whether the owl and the pussy cat
       went to sea.

   b. ?Where do you wonder whether the owl and the
       pussy cat went?

(2.98b) is not nearly as bad as (2.96bd), though it is hardly
a perfectly grammatical string. There doesn't seem to be a
foreground/background division in (2.98a), and so there is
no way of predicting the somewhat reduced grammaticality of
(2.98b). In this respect, then, (2.64) is no more adequate
than (2.25), and any revised island constraint ought to take
the data in (2.96) and (2.98) into account.

2.3.4.2 The rule in English of Left Dislocation is a poten-
tial counterexample to principle (2.64). It is a so-called
"copying" rule, that is, a chopping rule that leaves behind
a pronominal trace of the chopped element. Ross observed
(1967:236) that copying rules are not subject to his island
constraints.

Left Dislocation operates identically to the rule of
Topicalization, except it leaves behind a pronominal trace
of the fronted element, as illustrated in (2.99).
(2.99) a. Jonesy, I can't believe she wasn't in bed by six.

b. Those girls over there, did you notice that they are all hairless?

c. This milk (here), drink it!

For some speakers, but not for all, it is possible for the rule of Left Dislocation to remove elements with impunity from all sorts of restrictive environments: coordinate structures, relative clauses, cleft sentences, factive clauses, subordinate clauses, etc.

(2.100) a. Peter, Ivan and he never got along very well.

b. Ralph Mintz, I know a girl who actually likes him.

c. Lucy, it was in Paris that she lost her virginity.

d. Christians, it's deplorable that they condone violent acts.

e. Sheepdogs, she smells bad because she sleeps with four of them.

Left Dislocation appears to be a foregrounding operation par excellence, and as such the output it produces ought to be subject to (2.64). But the grammaticality of (2.100) shows that something is wrong.

I believe that (2.64) can be defended against this putative counterexample. Let me hasten to add, however, that the conclusions I shall draw about the nature of Left Dislocation will not affect the island constraint I shall propose in chapter three, which will allow (2.100) under a "normal" analysis of Left Dislocation as a copying rule.
There is some evidence that Left Dislocation may be an extra-grammatical operation that is not subject to constraints like (2.64). Here are my arguments:

(i) There appears to be a general constraint in English that prevents elements in yes/no questions and imperatives from being reordered.

(2.101) a. *Beans, did John eat?

   b. *This slave, flog!

But Left Dislocation is impervious to such a restriction.

(2.102) a. Beans, did John eat \{ \begin{align*}
   & \text{some} \\
   & \text{any} \\
   & \text{them}
\end{align*} \} ?

   b. This slave, flog her!

(ii) Left Dislocations are infinitely paraphrasable—putting quite a strain on the grammatical rules (and the base rules) if this process is to be treated in the usual way. (I have attested all of these forms—and I know that there are many more.)

(2.103) \{ \begin{align*}
   & \text{As for} \\
   & \text{As to} \\
   & \text{Speaking of} \\
   & \text{Now that you mention} \\
   & \text{Its funny about} \\
   & \text{You mentioned} \\
   & \text{Referring to} \\
   & \text{On the subject of} \\
   & . \\
   & . \\
\end{align*} \} \\

   \begin{align*}
   & \text{foul odors, I own a} \\
   & \text{sheepdog that mass} \\
   & \text{produces \#(them).}
\end{align*}

(iii) Consider these sentences:

(2.104) a. (As for) John and Mary, he kicked her out of his place last week.

   b. Bill, Sue and the turtle, he told her to get it out of his bed.
These appear to be left dislocations, but formulating the rule to reorder several discontinuous constituents strains the grammar. Notice that Left Dislocation cannot apply to its own output, so that explanation won't work.

(2.105) *John, Mary, he kicked her.

These three arguments cast serious doubt on the existence of a process of Left Dislocation as a normal rule of English.

I would like to suggest that what appears to be Left Dislocation is actually a gestural utterance whose purpose is to bind a pro-form referentially. I call such a process a deictifying operation. The reference of a pro-form to be subsequently used in the discourse is "pointed out as a gesture" in advance of the utterance, so when an appropriate pro-form appears, it will be deictic.

Left Dislocation is an overt linguistic way of deictifying, but in reality any kind of gesture will do. Thus if I walk into Zelda's house with my friend, upon observing a half dozen sheep dogs distributed hairily about the living-room I might whisper She smells bad because she sleeps with four of them, and coincident with the utterance I would gesture at the sheep dogs, in effect binding the pronoun them to the dogs.

Given this explanation, the situations cited in i-iii are not the serious problems they might be if Left Dislocation were to be an ordinary rule of the grammar. (i) is
obviously no problem. In (ii) we are merely asserting the
infinitude of gestural possibilities. (iii) is so simply
explained by my hypothesis, and would present such difficul-
ties if Left Dislocation were formulated as a grammatical
rule, that I consider it empirical support of the highest
quality.

As final, further evidence that Left-Dislocation is
extra-grammatical, we might note that the deictification
effected by it can be extremely sloppy. I don't believe this
much sloppiness is permitted by any true rule of the grammar.

(2.106) a. (As for) noxious odors, I own a sheepdog that
farts a lot.

b. Restaurants, I often go to Sushi bars, or
to Lawrey's.

c. Horses, I like big roans and grey mares.

Left Dislocation, then, belongs to the gestural compo-
nent of language use, and as such is not expected to be
constrained by principles such as (2.64).

2.3.4.3. The Dravidian language Kannada\textsuperscript{16} has two modes of
relative clause formation. One of these modes (the native
one), operates by deleting an identical nominal in the em-
bedded clause, much as in Thai, while changing the main verb
in the embedded clause into a participle. This mode of rela-
tive clause formation is subject to (2.25) and (2.64), as I
demonstrated in detail in Rođman (1973).

Kannada has another mode of relative clause formation
in which the embedded nominal to be relativized is neither
deleted nor pronominalized, but left entirely intact with the appropriate form of a relative pronoun (or perhaps "relative marker") inserted on its left. This rule is able to relativize elements of relative clauses and coordinate structures (cf. (2.107) below). Although (2.25) does not explicitly mention "rules of insertion" it is clear that Ross intended the island constraint to include such rules and to prevent them from "crossing island boundaries."

Sentences (2.107), then, are counterexamples not only to (2.64) but to (2.25) as well. The island constraint I propose in chapter three, however, will handle these cases in a natural and unexceptional way.

(2.107) a. gōpālanu sīteyu yāva citraṇavu tegeda
   Gopa    Sita     rel. picture drawn

   kalāvidānattu prītisuttālō ā
   artist     loves      definitizer

   citraṇavu kondanu
   picture      bought

   Gopa bought the picture which Sita loves
   the artist who drew (that picture).
(2.112) b. khurciya mattu yāva sōfāda
of the chair and rel. of the sofa
nāduve avaru mējannu iṭṭiddārō
between they table having placed
ā sōfā muridide
definitizer sofa is broken

The sofa which they have placed a table
between a chair and (that sofa) is broken.

2.3.4.4

2.3.4.4.0. There are other counter-examples to (2.64) in
other languages. For example, Mandarin permits Object Pre-
posing to operate on sentences in which an NP has undergone
constituent questioning (i.e., a sentence like Beans, who
likes? is permitted). Since these cases will also present difficulties for the conclusions reached in chapter three, I will discuss them there. Meanwhile we switch modes of criticism of (2.64). In sections 2.3.4.1-3 we discussed very specific types of cases that were not totally in accord with (2.64). In this section we shall look at a couple of more general objections to the principle.

2.3.4.4.1 All of the arguments regarding fuzzy data and relative grammaticality that we levelled against Ross's Island Constraint (2.25) are equally apropos as criticisms of (2.64). Principle (2.64) is a constraint that may (or may not) find a place in a discrete grammar. In a fuzzy grammar—which I claim is more adequate anyway—this principle would have to be fuzzified or abandoned.

As we shall see in chapter three, there are a number of factors that "conspire" to strengthen or weaken islands, of which the foregrounding/backgrounding distinction is but one, so an observationally adequate island constraint will at most incorporate (2.64) as a sub-principle.

2.3.4.4.2. We have already seen that one of the weaknesses of (2.64) is the difficulty of defining what a foreground or background is independently of the constraint, and when some part of a sentence qualifies for one such special status or the other.

The practical problem is intimately connected with another problem, namely, how to implement (2.64) formally.
Where in a transformational grammar does one represent such a basically semantic concept as "background"?

When I seriously proposed (2.64) as a universal principle of grammar in Rodman (1972a) and Rodman (1973), I had in mind the following very sketchy model of grammar:

**SYNTACTIC COMPONENT**  **SEMANTIC COMPONENT**

<table>
<thead>
<tr>
<th>Base Rules</th>
<th>semantic interpretation</th>
<th>Semantic Rep (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td></td>
<td>Shadow $T_1$</td>
</tr>
<tr>
<td>Intermediate Structure (1)</td>
<td></td>
<td>Semantic Rep (1)</td>
</tr>
<tr>
<td>$T_2$</td>
<td></td>
<td>Shadow $T_2$</td>
</tr>
<tr>
<td>Intermediate Structure (2)</td>
<td></td>
<td>Semantic Rep (2)</td>
</tr>
<tr>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_n$</td>
<td>Shadow $T_n$</td>
<td></td>
</tr>
<tr>
<td>Surface Structure</td>
<td>Semantic Rep (n)</td>
<td></td>
</tr>
</tbody>
</table>

This is basically the so-called "Extended Standard Theory" of Chomsky with a modification borrowed from Montague Grammar. In Montague Grammar every syntactic operation has a corresponding semantic operation to keep track of any changes in meaning that a syntactic operation might entail. There is a large body of evidence that suggests that all non-housekeeping transformations have the potential of changing the meaning of a sentence, especially when meaning is construed as something beyond logical entailment. It is not my intention to take up that matter here however.
Assuming this is the case, then corresponding to each transformation that maps a syntactic phrase marker into a syntactic phrase marker, there is a semantic Shadow Transformation that maps a semantic representation into a semantic representation.

If there is a need in the grammar for more than one component of semantic representation (e.g. one for representing logical relationships, one for grammatical relationships, one for relationships of relative prominence, etc.), then there will be Shadow Transformations for each level.

It is not my intention to defend this model. It is a multi-processing model rather than a serial processing model (as in Generative Semantics), and as such is perhaps more like actual mental processing—but both serial and parallel processing models may be so far from reality that no such argument could choose between them. The validity or non-validity of the model in no way negates the thesis at hand, which is concerned with fuzzy grammars, and more specifically, with fuzzy islands. Given such a model, however, I envision principle (2.64) as operating entirely within one of the semantic sub-components as a well-formedness condition on semantic representations. I presume a better understanding of semantics, and its relations with syntax, will lead to the ability to be more specific about the nature of the semantic component, its sub-components (if any), Shadow Transformations, and constraints on both form (semantic representation) and function (shadow transformations).
The difficulty with (2.64) in any of the standard models is that it is unclear as to exactly what it constrains. It is not a constraint on transformations—a good deal of its motivation arises from the fact that there seems to be a universal constraint which (2.64) encompasses barring the questioning of a constituent in one conjunct of a coordinate structure, yet there are languages like Thai in which grammatical transformations play no role in either the formation of constituent questions or in the formation of coordinate structures. Thus I would have to either claim that (2.64) applies on some level of semantic representation that is as yet not defined, or that it operates in some very vague and complex way as a well-formedness constraint on syntactic structures.

All these difficulties with (2.64) that I have discussed in section 2.3, including the last one, have convinced me that (2.64) must be abandoned in favor of a constraint that not only overcomes the observational inadequacies we have seen in this chapter, but is more easily adapted to one or another of the "standard" frameworks within which it is desirable to work. In chapter three we pursue this goal.
Footnotes
Chapter Two

1 See Postal (1971) for an explication of these rules and for arguments against their conflation into a single rule.

2 Postal (1971) defines the notion "essential variable" as opposed to non-essential variables. So-called "chopping rules" reorder elements over essential variables.

3 Actually Ross's definition of island also included phrase structure dominated by a sentential subject NP as defined in the Sentential Subject Constraint, and the phrase structure dominated by a left-branching node as defined in the Left Branch Condition on Pied Piping. All these constraints and conditions are eventually discussed in chapter three.

4 The stipulation "unidirectional" is because Pronominalization, considered a deletion rule by Ross, does not obey the island constraints. Only the fact that Pronominalization is bidirectional (it operates "backwards") distinguishes it from other deletion rules that do obey the constraints.

5 Note the constraint does not, in our analysis, restrict the rule of WH-pronoun deletion, which is an uncontrolled rule.

6 This last statement is perhaps slightly exaggerated. The CNPC, from whose statement islands are defined, refers to "noun phrases with lexical head nouns." It is not always clear what constitutes a lexical head noun in a given language. It is not even clear that the question can always be answered noncircularly.

7 The following analysis of pseudoclefts and cleft sentences in Thai is extremely tentative, or, as one critic has bluntly put it "quite fishy." Nothing significant hinges on this particular analysis, which I use for purposes of exposition, and until my understanding of Thai improves to where I can give a more adequate analysis.
There are problems for Akmajian's analysis—which is nonetheless the best one I know. Thus the following cleft sentence does not have a pseudocleft source:

It was very quickly that he ran to the store.

* {How the way that } he ran to the store was very quickly.

And the following pseudocleft sentence cannot be made into a cleft:

What John is is plain stupid.

* It's plain stupid that John is.

Finally, the important class of sentences that foreground a prepositional phrase by means of clefting cannot be derived by Akmajian's analysis:

It's to John that they were speaking.

* (The one) who they were speaking was to John.

George Bedell has pointed out to me that there are languages (e.g. Igbo) with cleft sentences but no pseudoclefts, and other languages (e.g. Japanese) with pseudoclefts but no clefts. These facts, then, cast serious doubt on the universality of Akmajian's analysis.

The Ozark dialect. I am indebted to Suzette Elgin for providing me with these data.

For now we are going to try to replace that part of (2.25) relevant to the CNPC and the CSC.

This deletion rule will in fact have to be ordered after the rule of Topicalization.

(i) * Her hand, the princess threw herself at the huntsman's feet because he kissed.

Topicalization is a very late rule—post shallow-structure according to Lskoff—so the deletion rule is too. Moreover, whether one considers the WH-preposing rules to be post-cyclic or not, the deletion rule will be post-cyclic since Topicalization is post-cyclic. But this rule can operate entirely within a non-root S, so the fact that it must be post-cyclic violates Chomsky's general condition given as (2.70). Thus there is some doubt as to whether the relative-clause analysis of adverbial subordinate clauses is viable.
Hooper (1973) has posited a class of "weak assertive" predicates (e.g. think, believe, suppose, expect, etc.). She hypothesizes, correctly it appears, that only this class of (non-factive) predicates allows raising (and neg-transportation (see below)).

How many times have any of us come away from a conversation with a certain item of knowledge which we communicate to someone else, who may ask something to the effect: "How do you know that—did so-and-so say it explicitly?" Often you are forced to answer that you don't know, that the information just "came out" of the conversation. I often have the experience, and once or twice I have traced the source of the information down to a presupposition on the part of the person from whom I gained the information.

This argument is also given in Schachter (1973), in the course of which Schachter makes the following statement:

It turns out that existential presuppositions are expressed by relative clause constructions just in those cases where the noun phrase that contains the relative clause is a 'referring' phrase, as in (66)a [as opposed to (67)a--data given at end of quote], and that it is essentially this characteristic of the noun phrase, rather than the fact that the phrase includes a relative clause, that entails an existential presupposition.

Schachter's data: (66) a. I'm looking for the unicorn that punctured my tires.
   b. Something punctured my tires.

(67) a. I'm looking for a man that has travelled faster than the speed of light.
   b. Someone has travelled faster than the speed of light.

The question is, what is a "referring" phrase? It is not a phrase with a definite determiner because (i) there are phrases with nondefinite determiners that are nonetheless referring phrases, such as this analogue to Schachter's (67)a, with the nonintensional verb "find" substituted for the intensional "look for": "I found a man that has travelled faster than the speed of light." The existence of such a man is indeed presupposed. And (ii) there are phrases with definite determiners that fail to be referring phrases, such as: the man that has travelled faster than the speed of light can name his own salary. If a referring phrase is
simply a referential noun phrase, as I suspect it must be, then the question is begged since a referential noun phrase is precisely a noun phrase in which the existence of such-and-such an entity is presupposed.

15 (2.97c) is ungrammatical relative to a reading that questions the reason John talked to the bartender:

(i) John told Mary that the bartender he had talked to the other night because he was lonely had a scarred cheek.

It is also ungrammatical relative to a reading that questions the reason for the bartender's scarred cheek:

(ii) John told Mary that the bartender he had talked to the other night had a scarred cheek because he was slashed by a customer who was irate over the presence of an olive in his sloe gin fizz.

16 Nadkarni (1970) is my source of information about this language.
CHAPTER THREE

3.1.0 Islands and island constraints have been subjects of much discussion in the literature since Ross opened a Pandora's box of linguistic facts in his 1967 thesis. All of this work\(^1\) suffers from a fundamental theoretical defect, viz., the assumption that the "unmarked" state of affairs is for there to be no islands; special "marked" configurations, however, are thought to be islands and subject to various island constraints. In other words, grammars are conceived in such a way that all islands have to be specifically mentioned and defined.

I wish to explore the hypothesis that grammars be formulated so that it is the "unmarked" case (if we may continue to use this Pragueian metaphor) that certain basic configurations form syntactic islands. Thus I propose (3.1) as a grammatical principle.\(^2\)

(3.1) All non-root S's and NP's are (fuzzy) islands.

When we say islands are fuzzy, we mean that islands have varying degrees of "islandness" or strength. Violations of island constraints when islands are weak lead to less ungrammatical strings than violations of island constraints when islands are strong. The need for the notion fuzzy island is dictated by the data we observe, some of it noted in section 2.2.3. A "fringe benefit" of taking islands to be fuzzy is the formally simple principle (3.1).

93

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This principle comprises a main theoretical point of this work. I am proposing a view totally polar to the "standard" approach to the subject in that I am claiming that rather than formulate the grammar to account for cases where S's and NP's behave like islands, we should formulate it to account for cases where S's and NP's do not behave like islands. In a discrete grammar this point of view seems incorrect because of the high grammaticality of (3.2).

(3.2) The boy who John believes Henry sold a book to...
But in a fuzzy grammar, where one's attention is forcibly diverted to grammatical phenomena that seem to affect the degree of strength of islands, it is possible to explain (3.2) within the broader generality of (3.1) by noting a conspiracy of island-weakening factors that come to a focal point in cases exemplified by (3.2).

An attempt to justify this view will be made in this chapter, as well as a proposal on how to formalize the notion of a fuzzy island. An island constraint compatible with fuzzy islands will be given in 3.2.

3.1.1 As I stated at the very end of section 2.2.4, and just above, it is to be expected that a thorough analysis of island phenomena that does not shirk from coping with fuzzy data will involve a number of (semi-)independent grammatical factors that affect the strength of an island, sometimes in a conspiracy-like fashion. Once we have formally defined what we mean by a fuzzy-island, to which this section is
devoted, we can then proceed to investigate island strengthen-
ing and island weakening factors.

The mathematical notion of a fuzzy set has already
been defined in chapter one, so it is an easy matter for us
to define the notion fuzzy island. For now let us define
islands on the level of syntactic deep structure, and let us
confine ourselves temporarily to $S$-islands:

(3.3) The set of nodes dominated by any $S_i$-node belongs
to a fuzzy set $I_i$, which is the fuzzy island
associated with $S_i$.

Thus to each node $M$ dominated by some $S_i$, there will
correspond a coefficient $m^M_i$, the degree of membership of $M$
in $I_i$. Since more than one $S$-node may dominate a given node
$M$, there may be several coefficients $m^M_i$, $m^M_j$, $m^M_k$, etc. associ-
ated with $M$.

Up to now we have been talking loosely and heuristical-
ly about "island-strength" or "degree of islandness." These
are useful terms to use intuitively, but the usage is impre-
cise. What we have is the membership strength of a node
relative to an island, a quality which, if it were quantified,
would be some (simple) function of the coefficient of mem-
bership of node with respect to the island. As I explained
in chapter one, the study of non-discrete phenomena is pri-
marily concerned with ordering relationships and not with
actual values of membership coefficients in fuzzy sets. The
mathematics of fuzzy sets is a formally (and intuitively)
satisfying way of expressing ordering relationships that we
hypothesize are in principle non-discrete. Thus we may state that node $M$ has stronger membership in island $I_i$ than node $N$, and without committing ourselves to specific values, capture the relevant fact by the formula $m^M_i > m^N_i$.

3.1.2 A small amount of careful thought about islands in a nondiscrete grammar will immediately yield two facts.

(i) Some islands are stronger than others.

To maintain formal precision we should really say something like: "Two nodes, corresponding in position and type, in two different islands may have differing degrees of membership in their respective islands." A linguistic universal may be stated at this point:

(3.4) If two nodes $M$ and $N$, corresponding in position and type, have differing coefficients of membership in two islands $I_i$ and $I_j$ respectively, such that $m^M_i > m^N_i$, then for any two nodes in the respective islands that correspond in position and type, say $M'$ and $N'$, it will be true that $m^{M'}_i > m^{N'}_j$.

We may illustrate this with a simple example: (3.5a) is less grammatical than (3.5b), a result that stems from the fact that the direct object NP letter has stronger membership in its island than the corresponding direct object NP note has in its island.

(3.5) a. The letter that I deplore that John wrote to his mother will be read tearfully.

b. The note that I am sorry that John wrote to his mother will be read tearfully.

The prediction that the universal principle (3.4) makes is that (3.6a) will be of lower grammaticality than (3.6b), which seems to be true.
(3.6) a. Who do you deplore that John wrote a letter to?

b. Who are you sorry that John wrote a letter to?

(ii) Within a single island, different nodes may have different degrees of membership.\footnote{5}

The principle example of this situation is that subjects have stronger memberships in a particular island than objects: \((3.7a) > (3.7b)\).

(3.7) a. The letter that I deplore that John wrote to his mother will be read tearfully.

b. The person whom I deplore that wrote a letter to his mother is named John.

3.1.3

3.1.3.0 In section 3.1.0 I stated that all S's are (fuzzy) islands. The qualification "fuzzy" is necessary because S's that are affected by none of the strengthening and/or weakening factors that we shall discuss here in 3.1.3 will turn out to be comparatively weak islands. Thus the complement of a non-factive predicate such as believe which is unaffected by any grammatical processes such as that-deletion, Raising or Extraposition will be so weak an island that objects will chop out with only the slightest lowering of grammaticality.

(3.8) What does the press believe that Ambassador Buggery argued with Colonel Deceit about?

The NP in the first PP, however, chops out a little less easily: \((3.9) < (3.8)\).

(3.9) Who does the press believe that Ambassador Buggery argued with about the presence of microphones in his toothpaste.
And subjects, as we've noted, chop out even less easily:

(3.10) Who does the press believe that argued with Colonel Deceit about the presence of microphones in his toothpaste.

(All these cases will be discussed in detail below.)

These, then, are the "neutral" cases. In the following sections we note how certain grammatical circumstances nudge the neutral, weak island in either a direction of strength, or of weakness.

3.1.3.1

3.1.3.1.1 I stated in section 3.1.2 (ii) that within a single island, different nodes may have differing degrees of membership. One of the hypotheses I wish to consider is that the degree of membership is dependent on horizontal position in the island. The more to the right (left), the smaller (greater) the degree of membership. Let me simply mention (with a promise of further discussion) what becomes immediately apparent: this factor acts to give subject-NP's the highest degree of membership in an island, subjects being more to the left than other constituents.

We must digress for two paragraphs to say something about phrase markers (Appendix I contains all the mathematical notions for making explicit the following remarks, and indeed, all remarks I make about relative positions of nodes in trees). I am assuming that phrase markers look something like the ones discussed in "Aspects." Chomsky's so-called "Extended Standard Theory" has modified things somewhat, but
not in a way that affects the relative horizontal ordering of nodes. On the other hand, underlying structures that many generative semanticists recommend may not preserve relative left-right relationships. Since I eventually propose that the membership coefficients of islands be updated after each derivational step, these differences will be of little consequence.

We have already seen evidence in several languages which supports that aspect of the hypothesis dealing with left-right relationships. In section 2.2.3 it was noted that it was more acceptable to question right-most constituents of conjoined NP structures in Korean, Ngizim and Thai. Moreover we gave data in the Ozark dialect of American English in which it was grammatical to question the right-most element, as in What did he eat peaches and?

We also showed in the same section that it was worse to question the subject of a relative clause in Mandarin than to question an object. In English, it seems to me to be somewhat worse to violate the CNPC by chopping a subject than by chopping an object—that is, (3.11b) < (3.12b)—though I admit both strings are so bad that a comparison may not be possible.

(3.11) a. I admire the ass on the girl whom your roommate is sleeping with.

b. Who do you admire the ass on the girl whom is sleeping with.
(3.12) a. I admire the ass on the girl who is sleeping with your roommate.

b. Who do you admire the ass on the girl who is sleeping with?

Perhaps a better example of the weaker island membership that more-to-the-right elements have is provided by (3.13), where (3.13b) > (3.13c).

(3.13) a. There is a student who donates several rusty bottle caps to his widowed mother's support each week.

b. His widowed mother's support, there is a student who donates several rusty bottle caps to each week.

c. Several rusty bottle caps, there is a student who donates to his widowed mother's support each week.

Another instance of this phenomenon occurs when various NP's in the sentences of (3.14) are preposed out of their respective prepositional phrases, stranding the preposition. For example, in questioning the NP $50,000, it seems that the farther to the right the prepositional phrase for $50,000 is, the better the result, despite the fact that there is more distance between the questioned NP and its stranded preposition when the prepositional phrase is rightmost. These results appear as (3.15): (3.15ab) < (3.15cd) < (3.15ef).

(3.14) a. Mr. Li claimed that he insured his house for $50,000 with the company against burglary.

b. Mr. Li claimed that he insured his house for $50,000 against burglary with the company.

c. Mr. Li claimed that he insured his house with the company for $50,000 against burglary.
d. Mr. Li claimed that he insured his house against burglary for $50,000 with the company.

e. Mr. Li claimed that he insured his house with the company against burglary for $50,000.

f. Mr. Li claimed that he insured his house against burglary with the company for $50,000.

(3.15) a. How much did Mr. Li claim that he insured his house for with the company against burglary?

b. How much did Mr. Li claim that he insured his house for against burglary with the company?

c. How much did Mr. Li claim that he insured his house with the company for against burglary?

d. How much did Mr. Li claim that he insured his house against burglary for with the company?

e. How much did Mr. Li claim that he insured his house with the company against burglary for?

f. How much did Mr. Li claim that he insured his house against burglary with the company for?

Grosu (1972:151f) provides two quite interesting instances of movement by a rightmost NP out of a coordinate structure. The first of these is in Hungarian: there may (or may not) be good reason for deriving (3.16c) from (3.16b), itself derived from (3.16a).

(3.16) a. a tanaar laatja a fiuut
    the teacher sees the boy

    ees ed' laaãt

    and a girl

    b. a tanaar a fiuut ees ed' laaãt laat

c. a tanaar a fiuut laatja ees ed' laaãt

In (3.16b) the rightmost element ees ed' laaãt is moved out of the coordinate structure to the right.
Samoan is the language of Grosu's other datum, which is attributed by him to Grinder. Apparently (3.17b) is derived from (3.17a) by a rule of topicalization.

(3.17) a. e sogi e le teine le tama i le to'i
cuts the girl the boy with the axe

    ma le naifi
    and the knife

The girl cuts the boy with the axe and the knife.

b. e sogi ma le naifi e le teine le tama
cuts and the knife the girl the boy

    i le to'i
    with the axe

Along with the knife, the girl cuts the boy with the axe (also).

Here we see that the rightmost element in the coordinate structure "the knife" has been moved to the left by a rule of Topicalization. In both this case and the Hungarian case, it is not possible to move non-rightmost elements in the same manner.

As I noted, subjects are left-most in S-islands and thus have strong membership in their islands. Data that we have already seen here and there in this work certainly bear out the claim that subjects are harder to question, cleft, relativize, etc. than other members in a given island.

In a language like English in which speakers seem to ascribe a role of extra importance to the subject, it is not surprising to find subjects as strongest island members. Not unrelated to this is the fact that in English the subject is
the only element (in unmarked word order) to precede the
predicate. This special relationship to the predicate may
contribute to the strength of the subject's island member-
ship.

In a VSO or SOV language that does not ascribe any
extra importance to the "subject," I would expect whatever
element constituted the "subject" (e.g. the agent of a causa-
tive verb) not to differ as significantly from other non-
subject elements with regard to its strength of membership
in an island as it does in English. This remains to be
proven, however.

We should note that it is not because English doesn't
allow subjectless clauses (for example, see (3.7b)) that sub-
jects might seem to have stronger island membership, for
(3.21) contains a subjectless clause and is certainly gram-
matical. Moreover, the that may or may not be present.

(3.21) She said (that) \{doing\} it on the head of a pin
would be impossible.

Even a tensed-S may be subjectless, though the that must not
be present. 8

(3.22) Who did John finally discover (?*that) broke
his Jew's harp?

To conclude this section we note that our hypothesis
on the varying degree of membership different nodes have in
islands is consistent with and supportive of Ross's hypothe-
ses on the notion of Primacy. Ross claims that node A has
primacy over node B just in case A is an element of a clause
above B, or, if A and B are clause mates, A is to the left of B. Although there are problems with the primacy constraint, Ross has assembled an impressive array of evidence that shows that languages recognize left-right distinctions in a regular, consistent manner. One manifestation of this state of affairs is, I have claimed in this section, degree of island membership.

(3.23) If A has primacy over B, then relative to a given island, A will have stronger membership than B.

3.1.3.1.1.2 Within a given island the degree of membership of a particular NP constituent position seems to vary slightly with the degree of identifiability of the referent of the NP. The identifiability squish is exemplified in (3.24).

(3.24) John is looking for some unicorn.
\{ a unicorn. \\
the unicorn. \\
that unicorn. \\
Mary's unicorn. \\
Mary's boyfriend's uncle's unicorn. \}

This hierarchy will play a significant role in section 3.3 where we discuss NP islands, and where we will see that the strength of certain NP islands is related to the degree of identifiability of the head noun.

The referent of "some unicorn" is lowest in identifiability. In fact the speaker chooses "some" rather than "a" to explicitly express his lack of knowledge or concern for the identifiability of what unicorn John is looking for.
"The unicorn" specifies the unicorn uniquely, but says no more, whereas "that unicorn" is accompanied by some additional identifying gesture. The referent of "Mary's unicorn" is uniquely specified with the added identifying fact of belonging to Mary. Finally the string of possessives strengthens the identity of the unicorn it modifies.

That degree of island membership is related to degree of identifiability is shown by (3.25), where (3.25a) ≤ (3.25b) ≤ (3.25c) ≤ (3.25d) ≤ (3.25e) ≤ (3.25f).

(3.25) a. Some unicorn, I hope that John is looking for.
   b. A unicorn, " "
   c. The unicorn, " "
   d. That unicorn, " "
   e. Mary's unicorn, " "
   f. Mary's boyfriend's uncle's unicorn, I hope that John is looking for.

3.1.3.1.2 We have just discussed the case where nodes within an island differ in their degree of membership. The rest of 3.1.3.1 is devoted to investigating factors that affect the overall strength of whole islands, without regard to particular nodes. In terms of our formal notions, when we say an island $I_i$ is increased in strength we mean that for each node $N$, $m_i^N$ is increased by some (multiplicative) factor.

A good part of chapter two was devoted to discussing principle (2.64), which in part stated that backgrounded
material had island-like properties. As I noted then, this was only one factor of many that had something to do with islands. We really want to assert something like the following:

(3.26) An island is strengthened if it happens to be backgrounded.

Thus backgrounding contributes to the strength of relative clause islands, clefted islands, and other island-structures discussed in chapter two, and for the reasons given there.

Backgrounding also plays a role in the strength of S-islands which are in apposition to abstract head nouns, as was noted in 2.3.3.2. At the conclusion of that section we noted, without explanation, a grammaticality cline of the type exemplified in (3.26-1): \( a > b > c \).

(3.26-1) a. Who did you make the claim that Sam was suspicious of?

b. Who did you believe the claim that Sam was suspicious of?

b. Who did you deplore the claim that Sam was suspicious of?

Many persons find (3.26-1a) as good as (3.26-2), with which it is synonymous.

(3.26-2) Who did you claim that Sam was suspicious of?

That synonymy led Ross (1967:77f) to attempt to derive (3.26-1a) from (3.26-2) by a rule of Modalization—an idea originally suggested by Harris. By perspicuously ordering this rule with respect to Relative Clause Formation, Ross hoped to account for why (3.26-1a) was not ungrammatical.
In Rodman (1971:114f) I argued that such a solution was impossible, and concluded that sentences like (3.26-1a) were counterexamples to the Complex NP Constraint.

While (3.26-1a) cannot be accounted for as we shall account for (3.26-1b) and (3.26-1c), there is nonetheless evidence that a solution is available. In constructs such as make the claim, make the proposal, make the statement, etc., the verb make is semantically empty and the constructs are synonymous with the simple verbs claim, propose, state, etc. Moreover such constructs behave syntactically as a single verbal unit. Thus the complementizer that, not normally deletable from an NP[N-S] construct, can be deleted from a make-NP[N-S] construct.

(3.26-3) Jeb is making the claim Mr. Mitchell is in it up to his eyebrows.

Cf.

? We are discussing the claim Mr. Mitchell is in it up to his eyebrows.

The entire make-NP construct can be gapped, just like the ordinary verb prefer, though this isn't true of ordinary NP[N-S] constructs.

(3.26-4) Helen preferred that Sven leave, and Anita that Gustaf stay.

Helen made the claim that the earth was going to explode, and Anita that the moon was going to explode.

??Helen discussed the claim that the earth was going to explode, and Anita that the moon was going to explode.
We note also that the determiner on the abstract head noun in the "make" constructs has constraints on it reminiscent of certain idioms.

\[(3.26-5)\] Myron is making \{ the his \} claim that the world \{ *Suzie's *their *my \} is a four-cusped hypocycloid.

Cf.

\[\text{Bill lost } \{ \text{the his } \} \text{ will to live.} \]

In light of Culicover (1967), where it is proposed that idioms are structured lexical items, I propose that make-NP constructs are also structured lexical items, as shown in \[(3.26-6)\].

\[(3.26-6)\]

\[
\begin{array}{c}
V \\
V_P \\
NP \\
\text{make} \\
\Delta
\end{array}
\]

Subcategorization rules of lexical insertion restrict the possible nouns to the appropriate class (i.e. abstract nouns that correspond to verbs such as state, propose, claim). Other theoretical devices necessary to this analysis are given in Culicover (1967) (e.g., make would be marked [+passive] to account for a proposal was made that-S, as opposed to kick [-passive] in *The bucket was kicked by Harry).

Of course the V in \[(3.26-6)\] is subcategorized to take an S-complement, which behaves like an ordinary S-complement.
to a verb such as claim with respect to its island forming capacity.

We must still consider why (3.26-1b) is better than (3.26-1c). When a speaker states I believe the claim that Sam was suspicious of Maurice, the object of that speaker's belief is principally the proposition expressed by the complement sentence Sam was suspicious of Maurice. The presence of the head noun claim is not particularly significant when the matrix verb is believe: the sentences in (3.26-7) are fairly similar in meaning.

(3.26-7) I believe the \{claim statement report\} that Sam was suspicious of Maurice.

In other words the head noun in these cases is not accorded augmented communicative importance. Thus the degree of grounding that the complement sentence is subject to is relatively small, and its island is therefore relatively weak.

In (3.26-1c), where deplore is the matrix verb, the situation is different. One may, for instance, deplore the claim that John is Jewish although he wouldn't deplore the fact or the possibility that John is Jewish. Thus deplore (or discuss, etc.) belongs to a class of verbs that causes the head noun of an NP[N-S] object to be foregrounded. The concomitant grounding increases the island strength of the S-complement.

If, in (3.26-1bc), we change the determiner on the head noun to the possessive Bill's, we force that head noun
to become higher in communicative importance (because it enters into a "possessive" relationship with another element) and hence add to the foreground/background dichotomy. That this is so is indicated by the fact that (3.26-8ab) are worse, respectively, than (3.26-1bc).  

(3.26-8) a. Who did you believe Bill's claim that Sam was suspicious of?  
b. Who did you deplore Bill's claim that Sam was suspicious of?  

Another situation analogous to what was just discussed revolves about the following data.  

(3.26-9) a. She's the kind of girl that there are many people who like.  
b. She's the kind of girl that I know many people who like.  
c. She's the kind of girl that we teased many people who liked.  

Despite the fact that in (3.26-9a) an element of a relative clause is itself relativized, the sentence is just slightly ungrammatical; (3.26-9b) is somewhat worse, and (3.26-9c) is terrible, although both these sentences seem to be parallel to (3.26-9a).  

The sequence (3.26-9c) is bad because girl is relativized out of the relative clause many people like the girl. As we have noted, the foreground/background dichotomy usually created by relative clauses is an important contributing factor to the strength of relative clause islands. In (3.26-9c) the outer relative clause is (3.26-10):  

(3.26-10) We teased many people who like the girl.
Many people is foregrounded; the modifying clause stipulating that the referent of many people is restricted to that set of persons who like a certain girl--the inner relative clause--is backgrounded. The semantic import of (3.26-10) revolves around the predicate tease, which relates the two noun phrases we and many people in a certain way. In addition, but of lowered semantic import, the noun phrase many people is qualified through a relative construction. When (3.26-10) itself becomes a relative clause, and girl is relativized out of it, a sequence of low grammaticality is created. This, of course, is what we would predict. What needs explanation are (3.26-9a) and (3.26-9b).

In (3.26-9a) the sequence analogous to (3.26-10) is given as (3.26-11).

(3.26-11) There are many people who like the girl. Because there are is semantically empty--the meaning of (3.26-11) is "many people like the girl"--the head of the relative construction many people does not receive the augmented communicative importance it would be entitled to if it were the object of a predicate that carried semanticity. That is, we do not view (3.26-11) as relating the noun phrase there to the noun phrase many people through the predicate be, as we did with tease in (3.26-10). (In fact *there are many people is not possible.) The sentence behaves semantically pretty much as if it actually were Many people like the girl. Thus the degree of backgrounding of many people
like the girl in (3.26-11)—which carries the entire meaning of the sentence—is naturally relatively low, and so the island is relatively weak. That is why (3.26-9a) is relatively high in grammaticality (but still not as high as She's the kind of girl who many people like, so a careful ordering of There-insertion is not the proper solution to this case).

The sequence (3.26-12) is a compromise between (3.26-10) and (3.26-11):

(3.26-12) I know many people who like the girl.

An uttering of (3.26-12) has a little more semantic content than merely uttering many people like the girl, but know does not carry the semantic import and novelty of a verb like tease when used in the first person, so the foreground/background division is less severe than in (3.26-10), though more severe than in (3.26-11). Another way of looking at it is this: if a person were to merely utter many people like the girl, it would be half expected of him to know those people, though it would be a complete surprise to discover that we teased those people (and no surprise to discover that there are such people). Note that if we change the subject of know from I to John in (3.26-9b) we get a sequence that is about as bad as (3.26-9c):

(3.26-13) She's the kind of girl that John knows many people who like.

This explication, which is difficult to express in concise terms, is valid I believe, and further supports my contention that (3.26) is a principle of natural language.
Another somewhat subtle situation that also suggests there are degrees of backgrounding can be seen by contrasting (3.27) with (3.28).

(3.27) The sentences that the linguist who tried to parse went insane are to be found in the appendix of my new book "Linguistic Pornography."

(3.28) The sentences that the man who tried to find went insane have never been uttered.

The sentence (3.27) is not quite as bad as (3.28). From a perceptual point of view the reason is clear: it's easier to associate "sentences" with "parse" in (3.27) than "sentences" with "find" in (3.28), since the range of possible objects of "parse" is considerably smaller than the range of possible objects of "find."

I claim that, in fact, (3.27) is less ungrammatical than (3.28). The reason is that the degree of the outermost foreground/background dichotomy in (3.27)--the one that foregrounds these sentences out of its background clause the linguist tried to parse these sentences--is reduced because of the close semantic association between the verb parse and the foregrounded object these sentences, with a resulting weakened background, hence weaker island, hence less of a violation of the island constraint. Needless to say the same conditions do not obtain in (3.28).

The island constraint to be formulated in section 3.2 will capture that part of principle (2.64) that forbids overlapping foreground/background distinctions, with the added advantage of mathematically precise fuzzification which

113
will capture subtle differences such as the ones described in this section.

3.1.3.1.3 In chapter two we noted that presupposed clauses were often backgrounded, and it was suggested there that perhaps presupposition entails backgrounding. It seems not unnatural for there to be some connection between these two phenomena, which both tend to give special status to clauses on the basis of a semantic property.

There are two types of presuppositional clauses that I wish to consider. The first of these, which I shall call incomplete presupposition, is relevant to WH-questions, relative clauses, and indirect question complements (Q-complements) and always involves a WH-word (in English). The second of these is the ordinary, or complete presupposition, usually associated with the complements of factive clauses. To my knowledge, this distinction has not been made before (though it is hinted at in Erteschik (1972)), but it is, I believe, significant.

Incomplete presuppositions are best characterized by indirect question complements.

(3.29) a. I didn't know who bought a book.

b. I didn't care how the Birdman escaped.

In (3.29a), the truth of the proposition expressed in the complement is presupposed, except the agent is unspecified: "Someone bought a book." In that sense the presupposition is incomplete. (In (3.29b), the presupposition is "The Birdman
escaped in some manner," and again it is incomplete insofar as the manner of escape is unspecified.

It is well known that indirect question clauses form fairly strong islands, and that suggests that we test the following hypothesis:

(3.30) Incomplete presupposition increases island strength.

This principle is one of the factors that conspires to make relative clause islands very strong. It also contributes to the island properties that WH-questions engender.

Indirect question complements provide the best examples of this principle at work. A careful study of island constraint violations involving Q-complements turns up the fact that they do not all give the same degree of ungrammaticality. There is a cline, exemplified in (3.31), where (3.31a) > (3.31b) > (3.31c).

(3.31) a. There's the book which I didn't know whether John bought for Mary.

b. There's the book which I didn't know who bought for Mary.

c. There's the book which I didn't know \{ why \} 

\{ how \} 

\{ where \} 

\{ when \} 

John bought for Mary.

Whether-clauses are the weakest kind of incomplete presuppositional clauses. What is incomplete, in fact, is the truth value of the proposition expressed by the clause. This type of clause is perhaps ill-labeled, since there is no presupposition of truth value, but rather a presupposition of "immediate relevance" (which accompanies all questions).
Since whether-clauses are at the bottom of the incomplete-preservation squish, we shall retain the terminology.

The string in (3.31b) is an example of a stronger incomplete-preservation in that the truth of the embedded clause, with a non-adverbial member specified, is presupposed: "Someone bought the book for Mary." The strings in (3.31c) represent the strongest kind of incomplete presupposition. The truth of the whole clause, with all non-adverbial members specified is presupposed, but an adverbial element is unspecified.

It appears that we can rank incomplete presuppositions from weakest (presupposition of "immediate relevance," e.g. whether Q-comps) to strongest (presupposition of entire clause with an adverbial element unspecified, e.g. why Q-comps). The data (3.31) suggest we modify (3.30):

\[(3.32) \text{ Incomplete-presuppositions increase island strength proportionately to the strength of the incomplete-presupposition.}\]

Note that the same relative grammaticality holds in all indirect question complements: (3.33a) > (3.33b) > (3.33c).

\[(3.33) \begin{align*}
\text{a. } \text{What is it } & \{ \text{unknown} & \text{a matter of} \} \text{ whether John bought for Mary.} \\
\text{b. } \text{What is it } & \{ \text{unknown} & \text{a matter of} \} \text{ who bought for Mary.} \\
\text{c. } \text{What is it } & \{ \text{unknown} & \text{a matter of} \} \text{ why John bought for Mary.}
\end{align*}\]

There is a potential counterexample to the incomplete-presupposition squish: While (3.32) predicts (3.34a) should be better than (3.34b), I don't find it so.

116

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(3.34) a. There's the book which I didn't know who John bought for.

b. There's the book which I didn't know why John bought for Mary.

Another principle is involved here, namely the one we discussed in section 2.3.1:

(3.35) Reordering more than one constituent from its slot in deep structure lowers grammaticality.

We noted certain problems with such a constraint that we weren't able to resolve satisfactorily, but it seems to have some merit as a perceptual constraint. This constraint behaves most vigorously when islands are involved and, significantly, when non-adverbial elements are reordered. It plays a significant role in the unexpected high degree of unacceptability we perceive in (3.34a) where two non-adverbial elements are displaced. In (3.34b) two elements are displaced, but one of them is the adverbial why. Adverbials in English are the most mobile type of element and speakers are accustomed, perceptually, to their displacement. Thus (3.35) should be modified to specify that when an adverbial is one (or more) of the reordered constituents, grammaticality is less drastically lowered. In (3.31b), of course, only a single constituent is reordered from its position in deep structure. We may conclude, then, that the principle (3.32), though on somewhat shaky ground, appears to have validity.

The inadequacy of a discrete grammar to account observationally for the facts of English is nowhere better
exemplified than in Chomsky (1971), where the following sentences would be "ungrammatical:"

(3.36) a. Is this the book which you asked me whether to buy?

    b. What beach were you wondering how to find?

    c. Is this the book which you asked me whether I ought to buy?

    d. What beach were you wondering how you ought to find?

A fuzzy grammar such as we are attempting to define would capture the fact that (3.36ab) > (3.36cd), respectively (see section 3.1.3.2 on island weakening processes for the reasons), as well as the fact that (3.36a) > (3.36b) and (3.36c) > (3.36d), leaving only (3.36b) and (3.36c) uncom-
pared.

It is somewhat of a mystery to me what the cognitive basis of (3.32) is. One is tempted to claim that speakers conceptualize presupposed material as possessing some inner cohesiveness that, naturally, has a strengthening effect on islands. If that were the case, and perhaps there is a half-truth hidden there, we would expect (non-incomplete) presupposed clauses to be most island-like. But they are not.

Though there is a tendency for factive clauses to form stronger islands than non-factive clauses, it is easy to find factive clauses (3.38) that are weak islands, and non-factive clauses (3.37) that are somewhat stronger islands, other things being equal.

118
(3.37) a. Who is it \{\text{definite}\} that John kicked?
   b. What is it \{\text{apparent}\}\{\text{improbable}\} that John ate?

(3.38) a. Who is it significant that John kicked?
   b. What are you sorry that John ate?

Nonetheless the conclusions drawn in 2.3.3.3, that presupposed clauses in general form stronger islands than non-presupposed clauses, seems worth preserving, if only as a first approximation, while we investigate the cases that are exceptional.

(3.39) Presupposition increases island strength.

There are three problems: (i) Why do certain non-presupposed clauses form stronger islands than expected (e.g. (3.37)); (ii) Why do certain presupposed clauses form weaker islands than expected (e.g. (3.38)); and (iii) Why do we have clines of grammaticality such as (2.54) in section 2.2.3, part of which we reproduce as (3.40).

(3.40) Whose hand are you sorry that the baby crapped in?
      Whose hand do you regret that the baby crapped in?
      Whose hand do you deplore that the baby crapped in?

The cline involving \textit{sorry}, \textit{regret}, and \textit{deplore} was first brought to my attention by Erteschik (1972), who explained the data by claiming that the "semantic complexity" of the matrix verb governed the strength of the island formed by its complement (to put things in our own terminology). Thus \textit{regret} contains the notion "sorry" plus something else.
(excessive extent of sorrow?), and is in this sense more complex. Similarly, deplore contains the notion regret plus something else, so it is more complex. Interestingly enough, Erteschik points out exactly the same cline of grammaticality in Danish. She also claims that infrequency of usage contributes to the semantic complexity of verbs. To further support her claim, Erteschik notes that the strength of a relative clause island is somewhat dependent on the "matrix" verb (meaning the verb with which the NP-relative clause is associated--this is not meant to be precise). Thus in

(3.41), (3.41a) > (3.41b) > (3.41c).

(3.41) a. This is the kind of weather that there are many people who like.

b. This is the kind of weather that I know many people who like.

c. This is the kind of weather that he teased many people who like.

Supposedly, be is less complex than know which is less complex than tease. Of course we saw in 3.1.3.1.2 a more plausible explanation for this particular case.

There seem to be other applications of Erteschik's observation. Quotative verbs (say, whisper, mumble, mutter, etc. etc.) display complements of varying island strength that does seem to depend on semantic complexity in Erteschik's sense. For example, complain is more complex than the unmarked verb say, since it means to say something with a particular attitude. Grumble is more complex than complain, for it means "to complain by muttering." The prediction, then, is for (3.42a) > (3.42b) > (3.42c).

120

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(3.42) a. Who did you say that John is dating?
   b. Who did you complain that John is dating?
   c. Who did you grumble that John is dating?

Similar results obtain between say and mutter, say and mumble, say and whisper, and so on.

Here's another putative example: (3.43a) \textgreater (3.43b).

(3.43) a. Who do you know that John is dating?
   b. Who are you positive that John is dating?

Be positive is more complex than know since it contains the meaning of know plus an extra assurance.

Erteschik's observations are insightful, but they contain a dangerous seed of tautology, since it is not frequently clear just which verbs are more "semantically complex" than which. For example, infuriated contains the notion angry plus something extra, so it is more complex. Yet it seems to me that neither string in (3.44) is worse than the other. (There are many pairs like this, so if the reader disagrees with me here and finds (3.44a) better than (3.44b), he can try it with some other pair.)

(3.44) a. Who were you angry that John kicked?
   b. Who were you infuriated that John kicked?

I see no way of systematically applying Erteschik's principle to (i) and (ii).

A study of this matter in sufficient detail to get a handle on what's happening is exceedingly difficult, perhaps impossible at this stage of our knowledge, simply because we
cannot be sure of our data with regard to strengths of various island-complements of various predicates. My own intuitions vary from day to day, if not from moment to moment. While tests can (and have been)\textsuperscript{13} devised to measure speaker's judgments of relative grammaticality, these tests are extremely difficult and tedious to carry out. A large body of informants is needed over an extended period of time. The data processing task is enormous, and all in all this type of research is only possible for those few with the available time and resources. On top of all this, as anyone can gather by reading the references given in footnote 13, the reliability of such tests is dubious, there being such a multitude of extraneous factors. One of the great drawbacks to studying fuzzy grammars is the difficulty of obtaining sufficient quantities of reliable data.\textsuperscript{14}

It nonetheless appears that Erteschik's solution puts us on the right path. The matrix predicate has some property (or properties) that affects the degree of strength of an island that is associated with it (as a complement, most commonly, but possibly as a relative clause object, etc.).

One candidate for such a property relevant to the complement islands of non-factive predicates may be "degree of belief." To the extent to which the speaker is asserting his belief in either the truth or falsehood of the complement, the complement shows a strengthened island. This is a sufficient condition, as (3.45) suggests, but probably not a
necessary condition (cf. the semi-grammatical (3.46)) according to my intuitions.

(3.45) ?Who is it \( \{ \text{true} \quad \text{definite} \} \) that John loves?
\( \{ \text{obvious} \quad \text{known} \} \) 
\( \{ \text{false} \quad \text{impossible} \} \) 
\( \{ \text{apparent} \} \)

?Who are you \( \{ \text{sure} \quad \text{positive} \} \) that John loves?
\( \{ \text{confident} \} \)

Counterexamples:

Who are you certain that John loves?
Who did you assert that John loves?

(3.46) ?Who is it \( \{ \text{uncertain} \quad \text{imaginable} \} \) that John loves?
\( \{ \text{predictable} \ \text{improbable} \} \)

?Who are you unsure that John loves?

?Who did you determine that John loves?

This partially explains why certain non-factives form stronger islands than we expect (relevant to (i)). The principle makes sense in that it seems plausible that a speaker tends to preserve intact a complement about which he has strong beliefs. A more adequate explanation of (ii) and (iii), however, than either Erteschik or I have been able to provide awaits further research.

3.1.3.1.4 In his forthcoming paper "Mouniness," Ross observes that the dichotomous distinction assumed to exist between the categories S and NP is actually non-discrete, and in order to capture many facts about the grammar of
English it is necessary to assign degrees of nouniness (or sentenceness) to structures. In particular, Ross gives a nouniness squish:

(3.53) \textit{that}-S (non-factive) \quad \text{We suppose that they resisted the police.}

\textit{that}-S (factive) \quad \text{We regret that they resisted the police.}

\textit{for} - \textit{to} \quad \text{We prayed for them to resist the police.}

\text{Embedded Q/headless rels.} \quad \text{We saw how bravely they resisted the police.}

\text{NP's V-ing} \quad \text{We adored their resisting the police.}

\text{Action nominal} \quad \text{We applauded their resisting of the police.}

\text{Derived nominal} \quad \text{We praised their resistance of the police.}

\text{Noun} \quad \text{We admire resistance.}

The hierarchy in (3.53) has the following property:

If some syntactic process affects the top member with a certain resulting degree of grammaticalness, then the application of that syntactic process to successively lower entries will result in successive degrees of grammaticalness that either increase all the way down, or decrease all the way down. Successive steps may result in no change in grammaticalness, but it will never be the case the grammaticalness first increases then decreases, or vice-versa, as you go down the hierarchy. In terms of fuzziness, consider the sequence of membership coefficients in the fuzzy set of English sentences that results as the syntactic process moves down the hierarchy.
Other things being equal, successive values in the sequence may be the same or greater for all members, or the same or smaller for all members, but values never increase and then a little later decrease, or vice-versa. For example in (3.54) a set of examples involving the rule of Extrapolation is given. The less sentence-like the less subject to the rule.

(3.54) a. It is likely that hopscotch causes nose bleeds.
   b. It is regrettable that too much sex activates the A-over-A principle.
   c. It was a pity for you to lose at four-square.
   d. It was ridiculous where you had to hide in order to listen to some Bruckner.
   e. It's crazy his winning the Nobel Prize for bootlicking.
   f. It's pleasurable the shooting of landlords.
   g. It's limitless my expectation of success.
   h. It's foolish Sally.

The reader may verify that \( a \geq b \geq c \geq d \geq e \geq f \geq g \geq h \).

Note that a counterexample to (3.53) would be a set of sentences like (3.54) where you couldn't state either \( a \geq b \geq c \geq d \geq e \geq f \geq g \geq h \) or \( a \leq b \leq c \leq d \leq e \leq f \leq g \leq h \), other things being equal (but see footnote 15).

In chapter four I note that the rule of Particle Movement interacts with (3.53). Particle Movement is a fuzzy-optional rule with respect to (3.53). A fuzzy-optional rule is a "normally" optional rule whose application, under certain circumstances, results in decreased grammaticalness. In
(3.55) we show that Particle Movement cannot take place over the less nouny entries, but as nouniness goes up, so does the grammaticality of moving a particle over the structure:

\[(3.55a) \leq (3.55b) \leq (3.55c) \leq (3.55d) \leq (3.55e) \leq (3.55f) \leq (3.55g)\]

(3.55) a. I want you to write that linguists are having trouble finding jobs down.

(I want you to write down that linguists are having trouble finding jobs.)

b. Would you please call for them to dispatch a cab up.

(Would you please call up for them to dispatch a cab.)

c. Let's look what they bought at the candy store over.

d. The group wanted to talk his explaining the problem over.

e. The coed brought her teacher's explaining of the problem up in class.

f. The clerk wrote their explanation of the problem down.

g. A teacher shouldn't put her students down.

Ross noted a number of other interactions of (3.53) with the grammar in "Nouniness," leaving no doubt that the squish is part of the competence of native speakers of English.

We move ahead, then, to the main result of this section:

(3.56) An increase in nouniness increases island strength.

This principle predicts that in (3.57), \( a \geq b \geq c \geq d \geq e \geq f \geq g \):
(3.57) a. Who do you believe that John gave a present to?
b. Who do you regret that John gave a present to?
c. Who did you plead for John to give a present to?
d. Who did you dislike John's giving a present to?
e. Who did you approve of the criticizing of?
f. Who did you approve of the criticism of?
g. Who did you approve of the father of?

I have omitted Q-complements because other factors are involved. Note that (3.58), which should go between (3.57c) and (3.57d), would be counterexemplary because it is worse than (3.57d).

(3.58) Who do you know why John gave a present to?

As we discussed in detail in 3.1.3.1.3, there are other factors involved in the conspiracy to make Q-complements strong islands. Their relatively high degree of nouniness is just one of those conspiring factors.

3.1.3.2

3.1.3.2.0 In this section we will talk about the one island weakening factor that I know about. It will encompass a variety of syntactic processes. The discussion is concerned with yet another non-discrete aspect of the grammar, viz., the notion of "clauseness." Clausal complements can occur in various states of collapsedness relative to the matrix clause—the higher the degree of collapsedness, the lower the degree of clauseness and the lower the degree of island strength.
3.1.3.2.1 A recalcitrant and frustrating problem has to do with the conditions under which the that of a that-S verbal complement may be deleted. As I concluded in Rodman (1972), the rule that deletes that is a fuzzy-optional rule insofar as its operation results in various degrees of grammaticality, from perfectly grammatical (3.59a) to fairly ungrammatical (3.59d).

(3.59) a. They thought the music was too loud.
b. They complained the music was too loud.
c. They regretted the music was too loud.
d. They rejoiced the music was too loud.

One of the difficult elements is trying to get grammaticality judgments. In preparation of my paper just cited, I had determined the grammaticality of a number of strings after that-deletion had operated. Later I ran across the paper "Some Classes of Verbs in English"--already discussed in footnote 14 of section 3.1.3.1.3--in which the authors indicated where they thought that-deletion was permitted. My judgments of grammaticality, and theirs, were in near random disagreement, leading me to conduct an elicitation experiment with eighty lower division students at UCLA that purportedly measured their collective intuitions, with a third set of results.

Despite the very fundamental difficulties (cf. footnote 14) involved in getting data, we may still achieve some insight into the workings of islands by a consideration of
that-deletion. One thing is certainly clear: deleting the that-complementizer of a complement clause weakens its status as an island—the sentences (3.60) are better than those of (3.61).

(3.60) a. Who do you believe ate beans on Tuesday?
   b. What do you believe John ate on Tuesday?
   c. On what day do you believe John ate beans?

(3.61) a. Who do you believe that ate beans on Tuesday?
   b. What do you believe that John ate on Tuesday?
   c. On what day do you believe that John ate beans?

Storing this fact away for a moment, let us digress to consider the relationship between a matrix sentence and a clause embedded in it. We have already seen that a "clause" may be more or less sentencelike/nounlike. Another insight due to Ross, suggested in his paper "Clause-matiness" (1972), is that the independence of different types of embedded clauses varies. In other words there is a squish—call it the clausiness squish—at one end of which the embedded clause is fully independent (highest degree of clausiness), and at the other end of which it collapses totally and is absorbed by the matrix (lowest degree of clausiness). There is some syntactic evidence in support of the necessity of such distinctions.

In "Clause-matiness" Ross gave a squish of nine structures with decreasing degrees of clausiness, of which we shall consider just the top five. They are, from most clausy to least clausy:

129
Factive that-clauses
Non-factive that-clauses
Tenseless that-clauses
Subjected non-finite clauses
Subjectless non-finite clauses.

The clausiest factive that-S clauses have no Verb-Verb restrictions, no sequence-of-tense restrictions, and little Raising:

(3.62) I regret( ted) that she \{will skyjack\} a Concorde. \{has skyjacked\} \{had skyjacked\}

?They \{knew\} acknowledged\{\} him to be a pimp.

The somewhat less clausy non-factive that-S clauses have no Verb-Verb restrictions, but their decreased independence is reflected by the frequent presence of sequence-of-tense restrictions and the much greater opportunity for Raising:

(3.63) He \{hoped\} \{claimed\} that you \{were\} \{are\} insane. \{believed\} \{thought\}

He \{believed\} \{expected\} the Sultan to be the father of 900 \{claimed\} \{hoped\} bloodthirsty religious fanatics.

Even less clausy are the tenseless that-clauses, which are thought to require one of several possible modals in deep structure, a Verb-Verb restriction. The modals are later deleted.

(3.64) He \{should\} \{must\}, we insist, be tickled with an ostrich feather.
We suggest that he \{ not be \} castrated.

Most verbs that take tenseless that-S complements also permit raising— in fact, it's often obligatory.

(3.65) He \{ asked \} \{ that she \} reveal the names of all the agents.

?He \{ prefers \} \{ wants \} that she bite her toenails before cocktails.

He \{ prefers \} \{ wants \} her to bite her toenails before cocktails.

**Subjected non-finites** have strong Verb-Verb restrictions—the embedded clause cannot be stative, and it generally denotes an event. Syntactically, the subject is always raised, of course. These are definite signs that such clauses are partially collapsed:

(3.66) I \{ saw \} \{ heard \} \{ owning a sheepdog. \} \{ running in place. \}

**Subjectless non-finites** require the like-subject constraint to hold in semantic representation. The embedded predicate must be volitional and must have an agentive semantic subject. Such complements are even low enough in clausiveness to permit reflexivization.

(3.67) Harry \{ wondered whether \} \{ learned \} \{ tried \} \{ began \} to start the engines.

Harry \{ ? learned \} \{ remembered \} \{ failed \} \{ didn't know how \} to start the engines (*by accident*).
Harry \{ asked Sue where \} managed \{ neglected \} \{ hesitated \} to teach himself how to swim.

These previous five paragraphs have provided some of the evidence in support of the existence of the clausiness squish.

We return now to our discussion of \textit{that-deletion}. The effect of \textit{that-deletion}, I claim, is to lower the clausiness of the embedded complement. That this is so is certainly intuitively plausible. The clauses are brought "physically" closer together. Moreover, in all but the clausiest complements, the \textit{that}-complementizer is absent. Raising, a clausiness reducing operation, also eradicates the complementizer.

There is some syntactic evidence in support of this claim, though it is not particularly compelling. Since complements of factive predicates are highest on the clausiness squish, we expect \textit{that-deletion} to be least allowed when the predicate is factive. By and large this is true, though there are exceptions in both directions.

(3.68) He resented *(that) she won.

He is sorry (that) she won.

(3.69) He believed (that) she won.

He disbelieved *(that) she won.

Another argument that unfortunately relies on a very delicate grammatical \textit{ity} judgment is as follows. The clausier a complement, the more independent of sequence-of-tense restrictions. \textit{That-deletion}, then, if it truly reduces

132
clausiness, should increase the requirement for sequence-of-tense restrictions. The sentences in (3.70) are equally grammatical. In (3.71), if the first sentence can be judged more grammatical than the second, then the argument holds.

(3.70) He understood that she \{was\} coming.

(3.71) He understood she \{was \} coming.

Finally, suggestive evidence that the deletion of that reduces clausiness is to be found by considering another one of Ross's works (1972). Ross gives the following category squish: (I've left out several irrelevant categories).

(3.72) Verb
   Present Participle
   Passive Participle
   Adjective
   Noun

There is some slight evidence that suggests that the lower, or nounier, entries in (3.72) take clausier complements. Ross (op. cit.) points out that "Raising applies 'more' to V than to A,..." and that "Raising is less applicable for N than for other categories..." As we have seen, the ability to raise indicates complements of lower clausiness.

In (3.73) we find that the that becomes increasingly reluctant to delete as we proceed from (a) to (e), suggesting once again the relationship between lower clausiness and the deletion of the complementizer: \( a \geq b \geq c \geq d \geq e \).
(3.73) a. He noticed the anaconda had swallowed a puppy.
   b. He was noticing the anaconda had swallowed a puppy.
   c. It was noticed the anaconda had swallowed a puppy.
   d. It was noticeable the anaconda had swallowed a puppy.
   e. The notice/notification the anaconda had swallowed a puppy...

At the beginning of this section we noticed that after that-deletion, the S-complement becomes a weaker island. Having established a link between that-deletion and degree of clauseness, we proceed to the main result of this section.

(3.74) When clauseness is reduced, island strength is reduced.

Any process (e.g., Raising, that-deletion, etc.) that decreases the clausal status of an S-complement, also reduces the strength of the island, and is hence an island weakening process. Note that (3.74) automatically accounts for the observation that that-deletion reduces island strength once we note that that-deletion reduces clauseness.

Further demonstration of principle (3.74) is found in the fact that (3.75a) ≤ (3.75b) ≤ (3.75c) ≤ (3.75d) ≤ (3.75e).

(3.75) a. Who do you resent that John laughed at?
   b. Who do you hope that John laughed at?
   c. Who do you suggest that John laugh at?
   d. Who did you hear John laugh at?
   e. Who did John try to laugh at?
Note also that we have accounted for the data in (3.36), as promised; (3.36ab), with less clausy complements, are better than their counterparts (3.36cd) when the clauses are relativized into.

(3.36) a. Is this the book which you asked me whether to buy?
   
b. What beach were you wondering how to find?
   
c. Is this the book which you asked me whether I ought to buy?
   
d. What beach were you wondering how you ought to find?

Also supportive of principle (3.74) is the fact that the complements to the category types in the category squish, which I argued become more and more clausy as you proceed down the hierarchy, also form successively stronger islands:

In (3.76), \( a \geq b \geq c \geq d \geq e \).

(3.76) a. Who do you think that George bit?
   
b. Who are you thinking that George bit?
   
c. Who is it thought that George bit?
   
d. Who is it thinkable that George bit?
   
e. Who does John hate the thought that George bit?

(I must add parenthetically that I have known for a long time that complements to passivized verbs and to adjectives form stronger islands than complements to corresponding verbs in the active (non-adjectival) mode. Here, finally, I believe, is a way to account for this fact.)

We note two interesting cases that the hypothesis put forward in this section explains. First, the strings in
(3.77ab) are better, respectively, than those in (3.78ab), where in both cases elements have been chopped out of adverbial clauses.

(3.77) a. Who did John make a fortune by cheating?
   b. What did John get a cramp while playing?

(3.78) a. Who did John make a fortune because he cheated?
   b. What did John get a cramp after he played?

Adverbial clauses normally form fairly strong islands (see section 3.2.2.3.2), accounting for the badness of (3.78). In (3.77), however, we have selected adverbial clauses that permit Equi-NP-Deletion to take place, and because this rule reduces clausiness, it is an island weakening process, and this accounts for why (3.77) are better than (3.78).

The second case is quite similar. Many nouns take for-to S-complements, and many take subjectless non-finite clause complements. Both are low in clausiness, with the latter lower than the former. This gives rise to the prediction that (3.79a) \(\leq\) (3.79b) \(\leq\) (3.79c).

(3.79) a. The horse that we rejected the \{decision\} suggestion
   that the company buy was beaten in the fourth race by a three-legged tapir.

   b. The horse that we rejected the \{decision\} suggestion
   for the company to buy was beaten in a race by a three-legged tapir.

   c. The horse that we rejected the \{decision\} suggestion
   to buy was beaten in a race by a three-legged tapir.
Sentences (3.79ab) would certainly be blocked by the CNPC, though (3.79b) does not seem all that bad. Depending on the status of pruning, (3.79c) may or may not be blocked by the CNPC. But the real facts, that there is a cline in (3.79), can only be captured in a fuzzy grammar, with help from the hypotheses set forth in this section.
3.2

3.2.0 In dealing with islands, there are two fundamental questions to consider. The first of these is what kinds of grammatical entities comprise an island, i.e., what are islands? We suggested in 3.1 that all non-root S's and NP's are islands, and that all islands are fuzzy. We confined our attention to S-islands in 3.1.3, where we studied a variety of island strengthening and weakening factors.

Having now a working hypothesis as to what comprises an island, we can turn our attention to the second fundamental question: what kind of constraints must be applied to islands? It is this question that we undertake to answer in 3.2.

3.2.1

3.2.1.0 Mathematical preliminaries.

Since ultimately the constraint I suggest be imposed on islands will be mathematically precise (although fuzzy concepts are involved) it is necessary to review some concepts of elementary set theory.

Consider two non-empty sets \( A = \{ a_1, a_2, \ldots, a_n \} \) and \( B = \{ b_1, b_2, \ldots, b_n \} \). We say that \( A \) contains \( B \) if and only if (iff) \( b \in B \implies b \in A \), where \( \implies \) denotes entailment. Similarly, of course, \( B \) contains \( A \) iff \( a \in A \implies a \in B \). We say that \( A \) and \( B \) are disjoint if their intersection is the null set \( \emptyset \), i.e., if \( A \cap B = \emptyset \). When two sets \( A \) and \( B \) are disjoint, it is always the case that \((a \in A \implies a \notin B) \& (b \in B \implies b \notin A))\) is true for all elements \( a \) and \( b \).
I wish to define the notion of overlapping. We say two sets overlap iff they are not disjoint, nor does one set contain the other. When two sets A and B overlap it must be the case that: (there is at least one \( a \in A \) such that \( a \notin B \))&(there is at least one \( b \in B \) such that \( b \notin A \))&(there exists some element \( x \) such that \( x \in A \)&(\( x \in B \))). Note that "overlap" is a set theoretically defined notion, just like "union" or "intersection"; the formal definition does, for the most part, correspond to the notion commonly denoted by the word overlap.

We are working (implicitly) in discrete set theory. Ultimately we must define these notions in fuzzy set theory because islands are associated with fuzzy sets, and we are interested in islands. Still, for the sake of comprehensibility, let me state a preliminary version of the constraint on islands at the onset, and illustrate it as if the notion island were a discrete one.

(3.80) Islands may not overlap.

This constraint reflects the fact that users of language regard islands as whole linguistic units. This can be done if the islands are disjoint, or if the islands are nested. But if the islands overlap, they lose their status of unity somewhat, and lowered grammaticality results.

3.2.1.1 Before we "fuzzify" (3.80), we must consider two crucial questions: (i) On what level does (3.80) take effect? and (ii) On what level(s) must islands be defined?
The answer to (i) is simple enough. Since Topicalization, a post-shallow-structure rule, is an island forming rule, (3.80) must be a surface structure constraint.\(^{18}\)

The answer to (ii) is not nearly so easy to arrive at. It does not even seem likely to me that islands are to be exclusively defined in the "syntactic component" (cf. 2.3.4.4.2). Because island strengths are affected by such semantic factors as presupposition and degree of backgrounding, it is fairly clear that semantic-level information must contribute to the determination of islands. On the other hand, islands are affected by those syntactic processes discussed in 3.1.3.2 (island weakening factors) and 3.1.3.1.1 (differences in island strength between nodes).

Since I know so little about what a semantic component might look like, or even what a semantic representation might look like, I shall attempt to give a principle for determining islands that operates in the syntactic component.

Since the strengths of islands are partially determined by their derivational history, a purely deep structure, shallow structure or surface structure principle would be inappropriate. As a first approximation we might consider this principle:

\[(3.81) \text{ All non-root S's and NP's are fuzzy islands in deep structure. After each derivational step, the status of existing islands is updated, and the introduction of new islands (by a rule such as Topicalization) is noted.} \]

That this principle is incompatible with (3.80) can be seen immediately by considering a simple example:
(3.82) *John, they arrested Harry because Max shot.

This underlying representation shows the island associated with $S_1$ ($S_1$-island) as comprising the elements \{Max, shot, John\}.\footnote{19} After the rule of Topicalization applies a derived structure something like (3.83) occurs:

(3.83)

$S_0$-island = \{they, arrested, Harry, because, Max, shot\}, and $S_1$-island is updated to \{Max, shot\}. The former set contains the latter, so (3.80) is not violated--an incorrect result.
The problem is immediately apparent: $S_1$-island should really have John in it, even though John is eventually proposed by Topicalization. The violation hinges on the proposing of John, so it is not surprising that our attention is drawn to this element. One is immediately tempted to have islands defined cyclically:

(3.84) Same as (3.81), except when the cycle containing the island is passed, no further changes to the island are made.

The previous case ((3.82) & (3.83)) will turn out all right, since $S_1$-island $= \{\text{Max, shot, John}\}$, and now the two islands will indeed overlap. But consider (3.85).

(3.85) *It was John who they arrested Harry because Max shot.

According to our revised principle, $S_2$-island $= \{\text{Max, shot, John}\}$ and since $S_1$ is not "broken up" until the $S_0$ cycle,
$S_1$-island = \{they, arrested, Harry, because, Max, shot, John\}. Again we have containment where we should have overlapping, and again we need to revise (3.81), this time as (3.86).

(3.86) Same as (3.81), except when the cycle that properly contains the island is passed, no further changes to the island are made.

Thus we stop updating $S_2$-island at the end of the $S_1$ cycle, and we stop updating $S_1$-island at the end of the $S_0$ cycle. In this case $S_1$-island = \{they, arrested, Harry, because, Max, shot\}, and $S_1$-island and $S_2$-island overlap in violation of (3.80)—a correct result.

We should back up to the example given as (3.82).

(3.82)

It is important to retain the notion (supported by Postal, Emonds, Chomsky, Lakoff and others) of post-cyclic rules, as discussed in footnote 20 above. If, for example, the $S_0$ cycle includes the rule Topicalization, John could be proposed out of $S_1$ before $S_1$ is "frozen" by the termination of the $S_0$-cycle, and there would be no overlapping islands in

143
*John, they arrested Harry because Max shot. The notion of
the post-cyclic rule has found its way into the theories of
many linguists, among whom I number myself, for (3.86) works
quite well once this distinction, well motivated on indepen-
dent grounds, is made.

We must make an additional general remark that pertains
to English (and to languages that relativize similar to
English). Consider this string:

(3.87) *Sue phoned the man who they arrested Harry
because Max shot.

\[
\begin{array}{c}
S_0 \\
NP \\
Sue \\
V \\
phoned \\
the \\
NP \\
man \\
they arrested Harry because \\
S_1 \\
\end{array}
\]

\[
\begin{array}{c}
Max shot the man \\
\end{array}
\]

Bearing in mind my analysis of relative clause formation
given in 2.1, and remembering that WH-preposing is post-
cyclic, we get the following results when we determine islands
according to (3.86): $S_2$-island = \{Max, shot, the man\}; and
at the end of the $S_0$ cycle $S_1$-island = \{they, arrested,
Harry, Max, shot, WH-man\}. For purposes of determining
overlapping we must state explicitly that man \# WH-man, and
then (3.80) makes the correct prediction.
This convention is a reflection of the fact that in English, a WH-word is generally associated with a foregrounded element. Since foregrounding is an island strengthening process (i.e., a background is concomitantly created), overlapping islands may form whenever the WH-word is in a non-root S. The overlap hinges on the WH-element, which is in a background when considered an element of the lower S, and not in a background when considered an element of the higher S.

There is one other point to consider. In certain dialects of English, Relative Clause Formation leaves a pronoun behind in the embedded clause.

(3.88) There is the girl who John likes her.

These dialects allow elements of subordinate clauses, relative clauses, and even coordinate clauses to be relativized. Let us consider (3.87) relativized in this way.

(3.89) Sue phoned the man who they arrested Harry because Max shot him.

$S_2$-island = \{Max, shot, the man$_1$\} and $S_1$-island = \{they, arrested, Harry, Max, shot, WH-man$_1$\}, just as in the ungrammatical (3.87). The problem here is that Pronominalization is a post-cyclic rule and, if we adhere to (3.86), cannot affect the content of any island. If Pronominalization were able to affect islands anyway—in contradiction to (3.86)—$S_2$-island would equal \{Max, shot, him$_1$\} and $S_1$-island would equal \{they, arrested, Harry, Max, shot, him$_1$\}. Now $S_1$-island contains $S_2$-island, and (3.80) does not lower the grammaticality of (3.89).
We must revise (3.86) to allow Pronominalization to affect islands passed by the cycle that properly contains them.

(3.90) All non-root S's and NP's are fuzzy islands in deep structure. After each derivational step, the status of existing islands is updated, and the introduction of new islands (e.g., by an island forming process such as Topicalization) is noted. After the completion of the cycle that properly contains an island, updating ceases for any derivational step but Pronominalization.

In addition, for English we need the convention:

(3.91) Two elements $X_i$ and WH-$X_i$ count as different for purposes of determining overlapping.

While the island constraint (3.80), or some version of it, is demonstrably universal (cf. Keenan, 1973), the principles by which means islands are determined is likely to contain language specific details. The non-English data that I give, however, is mainly to support or illustrate some particular point that is under discussion, or support the claim that (3.80) is universal, and I use the island determining principles on those languages that were worked out for English.

I think it is necessary to develop details in one language to show the feasibility of implementing the universal constraint against overlapping islands. English is a natural, but not a necessary, choice.

Analogously, while the island constraint (3.80) will be to an extent theory-independent—all theories will need to define islands, and once defined, some kind of constraint
on their interaction with the rest of the grammar will be necessary--the island determining principles are likely to be inextricably linked with whatever theoretical model we choose to adhere to. Thus in the Extended Standard Theory of Chomsky, where the cycle is defined on S's, NP's and AP's, principle (3.90) might have to be restated.

What is important about (3.90) is that it makes explicit the fact that islands cannot be defined on any one level. Also of importance is that the transformational cycle is needed in determining islands. This is because syntactic operations that take place a sufficient distance "above" an island cannot, except in well-defined circumstances, affect the status of the island.

We can make (3.90) formally precise:

(3.92) Let $M_i$ be any non-root S or NP in deep structure. For every node $N$ dominated by $M_i$ let there exist a coefficient of membership $n_i$ of $N$ in a fuzzy set $I_i$, the island associated with $M_i$. After each derivational step, each such $n_i$ is re-evaluated in accordance with the various principles set down in 3.1. If, in the course of a derivation, new S-nodes or NP-nodes are created, a corresponding island with its accompanying membership coefficients is also created. (However pruning of $M_i$ does not necessarily cause $I_i$ to cease to exist.)

After completion of the cycle that properly contains $M_i$, no $n_i$ may be changed except by Pronominalization.

3.2.1.2 One task now remains for us in this section, and that is to fuzzify (3.80)--the fundamental island constraint.

Consider two fuzzy sets $A = \{(a_1,n_1), (a_2,n_2), \ldots (a_N,n_N)\}$ and $B = \{(b_1,m_1), (b_2,m_2) \ldots (b_M,m_M)\}$. We say that
A contains B iff for all x the membership coefficient $m_x$ of x in B is less than or equal to the membership coefficient $n_x$ of x in A. That is $(x)(m_x \leq n_x)$. Conversely, B contains A iff $(x)(n_x \leq m_x)$. We say that A and B are disjoint iff it is the case that (for every x such that the membership coefficient $n_x$ of x in A is greater than zero, the membership coefficient $m_x$ of x in B is equal to zero) & (for every y such that the membership coefficient $m_y$ of y in B is greater than zero, the membership coefficient $n_y$ of y in A is equal to zero). That is, $A \cap B = \emptyset$ iff $(x)((n_x > 0) \rightarrow (m_x = 0)) & (y)((m_y > 0) \rightarrow (n_y = 0))$.

We must define the notion of overlapping in fuzzy set theory, and as in discrete set theory we may define it as the simultaneous absence of containment and disjointness: Two non-disjoint fuzzy sets A and B such that neither contains the other are said to overlap. That is, when two fuzzy sets A and B overlap it must be the case that: (there is at least one x such that its membership coefficient $n_x$ in A exceeds its membership coefficient $m_x$ in B) & (there is at least one y such that its membership coefficient $m_y$ in B exceeds its membership coefficient $n_y$ in A) & (there is at least one z whose membership coefficient $n_z$ in A is non-zero and whose membership coefficient $m_z$ in B is non-zero). Formally we have: A and B overlap iff the following is true: $(\exists x)(n_x > m_x) \& (\exists y)(m_y > n_y) \& (\exists z)((n_z > 0)\&(m_z > 0))$.

One of the crucial notions we deal with is strength of overlapping. In a discrete theory the closest we could come

148
to assigning degrees of strength to overlapping would be
to count up the elements involved in overlapping. Since the
highest state of overlapping lies somewhere between disjoint-
ness and containment, we would have to devise a method for
measuring the extent of a "partial intersection" of two sets,
and that parameter would be a measuring stick for the degree
of overlapping.\textsuperscript{23} The failure of such a scheme to capture
the myriad of facts being discussed in this thesis is self-
evident, for it certainly matters which elements overlap,
and to what extent any individual element overlaps—-notions
that belong ultimately to a non-discrete theory.

We shall find it useful in dealing with degrees of
overlapping to consider overlapping as comprising two distinct
sub-overlapping. That is, if two sets A and B overlap,
then we can talk about the degree to which A overlaps B and
the degree to which B overlaps A. The degree to which A
overlaps B depends on two things.\textsuperscript{24} (i) the membership co-
efficients of those elements of A that do not belong to B
and (ii) the difference between membership coefficients of
elements shared by A and B, where the membership coefficient
with respect to A exceeds that of B. The degree to which B
overlaps A is defined conversely. A concrete example will
suffice to clarify these notions. Let $A = (a_1, 0.8), (a_2, 0.2),
(x, 0.5), (y, 0.7)$ and $B = (b_1, 0.5), (b_2, 0.6), (x, 0.3),
(y, 0.9)$. A overlaps B to the extent that two elements
($a_1$ and $a_2$) belong to it that do not belong to B, and the
element x belongs to it "more strongly" than it belongs to B. The degree of overlapping is possibly proportionate to the sum of the membership coefficients of $a_1$ and $a_2$, and the sum of the difference between the membership coefficients of x in A and x in B. That would be $(0.8) + (0.2) + (0.5 - 0.3) = 1.2$. Similarly, B overlaps A by virtue of $b_1$ and $b_2$, which belong to B but not to A, and by virtue of the element y's stronger membership in B. The degree to which B overlaps A, according to the additive method is proportionate to $(0.5) + (0.6) + (0.9 - 0.7) = 1.3$. The total degree of overlapping is, if we persist in adding, proportionate to $1.2 + 1.3 = 2.5$.

The reason we are interested in considering overlapping as the combination of two asymmetric overlappings is that when we deal with islands, we will find that the degree of overlapping in one direction contributes much less heavily to ungrammaticality than the degree of overlapping in the other direction.

Some illustration from actual linguistic situations will be useful here. We shall, as usual, lapse into dealing with ordering relationships rather than with actual values of membership coefficients in fuzzy sets.

Consider the very weak island formed by the complement of the verb believe:

(3.93) a. I believe that John eats beans. (n)
    b. Its beans that I believe that John eats. (m)
    c. Its John who I believe that eats beans. (m)
The lower island is \{John, eats, beans\}, and \(n(\text{beans}) < n(\text{John})\) in accordance with the principle that "lefter" nodes have stronger membership in their island (cf. 3.1.3.1.1.1).

The higher island in (3.93b) is \{I, believe, John, eats\}. The lower island overlaps the higher island by a factor of \(n(\text{beans})\). The higher overlaps the lower by a factor of \(m(I) + m(\text{believe})\), assuming that \(m(\text{John}) = n(\text{John})\) and so on.

The degree of overlap in (3.93b) = \(D_b = n(\text{beans}) + m(I) + m(\text{believe})\). The degree of overlap in (3.93c), by the same token, = \(D_c = n(\text{John}) + m(I) + m(\text{believe})\). \(D_c > D_b\) because \(n(\text{John}) > n(\text{beans})\). This is the correct result since (3.93c) < (3.93b).

Now consider (3.93de):

(3.93) d. I resent that John eats beans. (p)

   e. It's beans that I resent that John eats. (m)

We have claimed that complements to factive verbs form stronger islands than complements to non-factive verbs, other things being equal. That is \(p(\text{beans}) > n(\text{beans})\). The degree of overlapping in (3.93e) = \(D_c = p(\text{beans}) + m(I) + m(\text{believe})\) and clearly \(D_c > D_b\) because \(p(\text{beans}) > n(\text{beans})\). This is the predicted result since (3.93e) < (3.93b).

The degree of ungrammaticality that results from a violation of the island constraint (3.80) depends on the degree of overlapping, and it appears from the examples we have just looked at that it depends more crucially on the contribution to total overlapping made by the lower island's

151
overlapping of the upper than vice versa, for we have seen two instances where differences in grammaticality are directly reflected in differences in the lower island's overlapping.

This same fact can be seen in a different way. Consider (3.94).

(3.94) It's beans that I believe that Max observed that John eats.

There are three islands to consider: The cleft or "higher" island: {I, believe, Max, observed, John, eats}; the complement of believe: {Max, observed, John, eats, beans}; and the complement of observe: {John, eats, beans}. Both lower islands overlap with the clefted one, suggesting a much greater violation than actually exists. In fact (3.94) is only slightly worse than (3.93b), a reflection of the principle that the farther a rippee has to travel from its point of departure, the worse the result. This slight difference can more adequately be captured if we adopt the following convention:

(3.95) If two islands A and B, such that A contains B, both overlap with a third island C, the constraint (3.80) considers only the degree of overlapping of the contained island B with C.

Now the contribution to overlapping is n(beans) (the lower overlapping the upper) and m(I) + m(believe) + m(Max) + m.observe (the upper overlapping the lower). This degree of overlapping exceeds the degree of overlapping we observed in connection with (3.93b), a proper result. The trouble is that it exceeds it by a factor of m(Max) + m(believe)---
somewhat out of proportion to the minimal difference in grammaticality between (3.94) and (3.93b). I have no precise solution to offer here. The point is that the grammar will have to discount the upper island's overlapping in computing the actual lowering of grammaticality.

There is some slight independent evidence in support of convention (3.95). Consider the sentences (3.96) and (3.97).

(3.96) It's the guard whom I deplore that John believes that Mark murdered.

(3.97) It's the guard whom I believe that John deplores that Mark murdered.

In (3.96) there are three islands to consider: the complement of believe: \{Mark, murdered, the guard\} (n); the complement of deplore: \{John, believes, Mark, murdered, the guard\} (m); and the cleft island: \{I, deplore, John, believes, Mark, murdered\} (p). There are three ways of determining degree of overlapping: (i) add up all overlapping: $n(\text{guard}) + p(I) + p(\text{deplore}) + p(\text{John}) + p(\text{believe}) + m(\text{guard}) + p(I) + p(\text{deplore})$; (ii) ignore lower island: $m(\text{guard}) + p(I) + p(\text{deplore})$; (iii) ignore upper island in accordance with (3.95): $n(\text{guard}) + p(I) + p(\text{deplore}) + p(\text{John}) + p(\text{believe})$. Now let's do the same thing for (3.97): the three islands are: the complement of deplore: \{Mark, murdered, the guard\} (n'); the complement of believe: \{John, deplores, Mark, murdered, the guard\} (m'); and the cleft island: \{I, believe, John, deplore, Mark, murdered\} (p').

153
The three ways of determining overlapping give: (i') add up all overlapping: \( n'(\text{guard}) + p'(\text{I}) + p'(\text{believe}) + p'(\text{John}) + p'(\text{deplore}) + m'(\text{guard}) + p'(\text{I}) + p'(\text{believe}) \);

(ii') ignore lower island: \( m'(\text{guard}) + p'(\text{I}) + p'(\text{believe}) \);

(iii') ignore upper island in accordance with (3.95):
\[
\text{If (i'), the degree of overlapping computes out to be greater for (3.96) than for (3.97). In cases (ii), since } \text{m(guard)} > m'(\text{guard}) \text{ (because m(guard) is the coefficient of membership in the factive complement of deplore), the degree of overlapping of (3.96) exceeds the degree of overlapping of (3.97). In cases (iii), since } n'(\text{guard}) > n(\text{guard}), \text{ the degree of overlapping of (3.97) exceeds the degree of overlapping of (3.96).}
\]

Now all we have to do is decide the relative grammaticality of (3.96) and (3.97) and this will provide a basis for choosing among (i), (ii) and (iii). If (3.96) < (3.97), we opt for (i) or (ii) and if (3.97) < (3.96) we opt for (iii). My (prejudiced) intuitions find (3.97), with the factive clause more deeply embedded, somewhat worse than (3.96), thus supporting (iii) and the principle given in (3.95).²⁵

We may now state the island constraint (3.80) in terms of fuzzy concepts and relative grammaticality.

(3.98) Given two strings X and Y, if the degree of island overlapping in X exceeds the degree of island overlapping in Y, then, other things being equal, \( X < Y \). Moreover, the degree of grammaticality of any string is (in part) a function (not known precisely) of the degree of overlapping of any islands it contains.
In this section we have defined in precise mathematical terms the theoretical concepts we have drawn on, and shall draw on for the remainder of this dissertation. For the most part we shall fall back into a more relaxed mode of rhetoric on the assumption that having once shown the possibility of mathematical precision, it becomes more important to give a comprehensible, if somewhat formally loose, presentation.

3.2.2

3.2.2.0 In this section I shall consider a wide range of cases and show that they can all be dealt with successfully by applying the principles discussed in 3.2.1.

3.2.2.1 In Ross (1967) and Rodman (1973), island constraints were given that could only exclude copying rules from their dominion by an ad hoc statement. By applying the principles (3.98) and (3.90), it follows that copying rules do not, in fact, result in violations of the island constraint.

(3.99) Ralph Mintz, I actually met someone who likes him.
At surface structure, where (3.98) takes effect, principle (3.90) gives the following results: $S_1$-island = \{who, likes, him\}; $S_0$-island = \{I, ... someone, who, likes, him\}. Since $S_0$-island contains $S_1$-island, there is no violation and (3.99) is predictably grammatical.

We have already seen that when relativization is a copying rule, the principles (3.98) and (3.90), in conjunction with convention (3.91), predict the correct results.

Note that we can now simplify the grammar by letting Topicalization and Left Dislocation be the same rule, with the option of leaving behind a pronominal trace of the topicalized element. The island constraint will lower the grammaticality of the rule's output in just the right cases.26

It is important to realize that if the island constraint is a constraint on rules, rather than on structures as I have defined it, this simplification is impossible without the ad hoc device of letting the constraint block certain options of the rule. Other rules (e.g. Relative Clause Formation in certain dialects of English) can be conflated this way with no further complication to the grammar.

These principles also work in a language like Kannada, discussed at some length at the end of 2.3.4.2. Kannada, it will be recalled, has two modes of relative clause formation, one of which deletes the embedded, relativized nominal (like Thai), and the other of which leaves the nominal intact, but with a morphological marker attached. This latter mode of
relative clause formation allows elements of relative clauses to be themselves relativized, a result we now show is predicted by the principles we are discussing. Consider (3.100), and an accompanying underlying structure.

(3.100) gōpālanu sīteyu yāva citravannu
gopal Sita rel. pro. picture
tegeda kalāvidanannu prītisuttālō ā
drawn artist loves definitivizer
citravannu konḍanu
picture bought

*Gopal bought the picture which Sita loves the artist who drew [that picture].

At the end of the $S_1$ cycle, the $kalāvidanannu$ of $S_2$ has been deleted by relative clause formation mode 1, leaving

157
$S_2$-island = \{citravannu, tegeda\}. At the end of $S_0$, Relative Clause Formation mode 2 has attached the marker ṭāva to citravannu, so $S_1$-island = \{Śiteyu, ṭāva, citravannu, tegeda, kalāvidanannu, prītisuttālō\}, and contains $S_2$-island. Thus (3.98) does not lower the grammaticality of (3.100).

It is important to note that when relativization is a "pronoun-leaving" rule in a language, the principles (3.98) and (3.90) predict that it will be able to "go into" a much greater diversity of structures than if it were non-pronoun leaving.

In fact, the formulations I have given so far predict that pronoun-leaving rules can penetrate anywhere, assuming they operate analogously to the pronoun-leaving rules such as English Left-Dislocation. This is a false prediction. Keenan (1973) has noted several languages that have pronoun-leaving rules of relativization that cannot penetrate everywhere. I myself have never studied a particular instance where such a relativization rule could not penetrate, say, an indirect question, as Keenan claims for Batak (Sumatra) and Urhobo (Nigeria). One must consider each language individually to see if other factors play a role besides the island constraint. There is much we don't know about transformational rules that disrupt island structure, irrespective of whether there is ultimately overlapping.

Having noted the worst aspects of my principles, I now hasten to reproduce the totality of results of Keenan just

158
alluded to, which strongly support the principles. In this table, reproduced from Keenan's handout (1973), it can be seen that in general relativization in pronoun leaving languages have much more penetrating power than relativization in non-pronoun-leaving languages.

Relativization is possible into:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aoba (New Hebrides)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no?</td>
<td>yes</td>
</tr>
<tr>
<td>Arabic</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Batak (Sumatra)</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Gilbertese (Gilbert Is.)</td>
<td>yes</td>
<td>no</td>
<td>yes?</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Hebrew</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Persian</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Urhobo (Nigeria)</td>
<td>yes</td>
<td>-</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Welsh</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Zurich German</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>yes</td>
<td>no</td>
<td>yes?</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Russian</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Malagasy</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Finnish</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Tagalog</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

There are two types of exceptions that I shall note, but not discuss since I am not acquainted with the languages involved.

159
First there are the pronoun-leaving languages that fail to penetrate where my principles predict they should. These correspond to the "no's" in the "+PRO" section. I would want to study each exceptional language in an attempt to find other reasons for the ungrammaticality (and, if possible, the degree or extent of ungrammaticality). If none are found, some doubt is cast on the universality of the principles as I have formulated them.

Secondly are the "no's" below the line in the columns of weaker islands, such as VP-S. These indicate that some languages ascribe higher degrees of island strength to a given structure than others or, alternatively, some languages are more sensitive to small degrees of overlapping than others. What is crucial here is that any language that says "no" to VP-S will say "no" to NP-S or Rel. Cl.--the latter containing stronger islands. This is borne out by the data.

3.2.2.2.1 Conjoined sentences form extremely strong islands in every language with which I am familiar. Just as the effect of a foreground/background semantic division strengthens islands (cf. 3.1.3.1.2), the semantic division of material into equal parts, as under coordination, also strengthens islands. This fact was noted in a more narrow context in Rodman (1973), where I observed that "once a division is made into equal parts, no element of any conjunct, nor any conjunct itself may be foregrounded at the expense of an equal partner."

160
We should demonstrate that the constraint (3.98), in conjunction with principle (3.90), operates correctly when conjoined clauses are involved:

(3.101) a. *I don't know where this little piggy went and that little piggy stayed at home.

b. *The home that this little piggy went to the market and that little piggy stayed at is insured against huffing and puffing.

Here is a tree diagram of a structure underlying (3.101a):

```
  S₀
  /  \
 VP--NP
     I--V
       /  \
  don't--know--S₁
           /  \
  this--little--piggy--S₂
               /  \         /  \
            went--somewhere--that--little--piggy
                                stayed--at--home
```

At the end of the \( S_1 \) cycle, \( S_2 \)-island = \{this little piggy, went, somewhere\}. At the end of the \( S_0 \) cycle, \( S_1 \)-island = \{this little piggy, went, WH-somewhere, \( S_3 \)\}. Since WH-somewhere and somewhere are considered different elements by convention (3.91), there are overlapping islands, and (3.98) lowers the grammaticality of (3.101a).

We are able to achieve the same results in case (b). Again, a tree diagram will facilitate the explication:
At the end of the $S_1$ cycle, $S_3$-island = \{that little piggy, stayed, home\}; at the end of the $S_0$ cycle, $S_1$-island = \{$S_2$, that little piggy, stayed, WH-home\}, and the islands overlap.

In languages that effect relative clause formation by deletion (as, in fact, English does optionally in some cases), the constraint continues to work, as this example in Thai shows:

(3.102) *khāw thîː maeːw dûːm nom lēː māː kin
rice comp cat drink milk and dog eat
phêt
spicy

The rice that the cat drank milk and the dog ate was spicy.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
At the end of the $S_1$ cycle, $S_3$-island = \{mā:, kin, khâw\};
at the end of the $S_0$ cycle, $S_1$-island = \{$S_2$, mā, kin\}, and
the islands overlap.

Similar results can be obtained when we regard structures of English that do not involve WH-preposing. We illustrate this fact with one such structure (but see Ross (1967: ch6) for several more). Consider:

(3.103) a. Maxwell isn't half the doctor his father was.

b. * Maxwell isn't half the doctor his uncle was a cop and his father was.

(a)

```
  SO
   /
  /  \\
 NP  VP

Maxwell

  /
 isn't half NP

  /
 S1

\triangle

Maxwell's father
was a doctor
```

There are clearly no overlapping islands.

(b)

```
  SO
   /
  /  \\
 NP  VP

Maxwell

  /
 isn't half NP

  /
 S1

\triangle

S2

  /

\triangle

S3

Maxwell's uncle
was a cop

Maxwell's father
was a doctor
```

163
In this case $S_3$-island = \{his father, was, a doctor\} and $S_1$-island = \{S_2, his father, was\}. The two islands overlap, and (3.103b) is predictably ungrammatical.

There is a major problem, however. According to my limited knowledge, and Edward Keenan's more extensive knowledge, no language ever allows relativization into any one of two or more conjoined S's, even if the rule is a pronoun-leaving rule. The principles we have been applying in this section predict incorrectly that relativization should be allowed in the pronoun-leaving languages.

This fact vitiates my principles somewhat, although no island constraints with which I am familiar (Ross's CSC, or Grosu's (1973) modification of it) fare any better.

Consider now the perfectly grammatical sentences of (3.104).

(3.104) Marc drives a gold Eldo and Rita drives Marc crazy.

John loves beans but Mary won't bring beans into the house.

The coordinate islands in (3.104) overlap because they share a common element. Principle (3.98) would lower the grammaticality of (3.104), a result we must avoid.

One solution for these particular cases is that we adopt the following convention:

(3.105) Different token occurrences of the same element count as different for purposes of determining overlapping.

164
There is plenty of independent motivation for such a convention. In none of the following sentences do we want overlapping islands.

(3.106) a. \( g[\text{John couldn't bring Sally over to meet Sue}], \) for \( g[\text{Sally hates her}]. \)

b. \( g[\text{That John could introduce Sally to Sue}] \)
\( g[\text{proves that Sally doesn't hate her}]. \)

c. \( g[\text{That John didn't dare introduce Sally to Sue}]; g[\text{whom Sally hates}]. \)

Thus we add (3.105) to the conventions and principles being investigated in this thesis.

3.2.2.2 In Ross (1967:98f) it was noticed that certain rules including Conjunction Reduction and Relative Clause Formation were capable of removing elements from coordinate structures in violation of the Coordinate Structure Constraint if they removed co-extensional elements from each and every conjunct. For example, Conjunction Reduction applied to

(3.107a) gives (3.107b):

(3.107) a. God decreed that Tom would pick these grapes and Sue would wash these grapes and I would prepare these grapes.

b. God decreed that Tom would pick, Sue would wash, and I would prepare, these grapes.
(3.107a)

```
(3.107a)  
(3.107a)  
S_0  
|   |   |
NP  VP  
|    |   |
God V  
|   |   |
S_1  
|   |   |
decreed S_2  S_3  S_4  
|   |   |   |   |
Tom would pick these grapes S_2  Sue would wash these grapes S_3  I would prepare these grapes S_4  
```

CONJUNCTION REDUCTION

```
S_0  
|   |   |
NP  VP  
|    |   |
God V  
|   |   |
S  
|   |   |
decreed S_1  
|   |   |
S_2  S_3  S_4  
|   |   |   |
Tom would pick S_2  Sue would wash S_3  I would prepare S_4  
these grapes  
```

Conjunction Reduction would apply on the S_1 cycle, so that S_2-island = \{Tom, would, pick\}, S_3-island = \{Sue, would, wash\}, S_4-island = \{I, would, prepare\}, S_1-island = \{S_2-island, S_3-island, S_4-island\}. None of these islands overlap so this exception to the CSC vanishes under the present analysis.

166
In arguing that Relative Clause Formation must operate across-the-board, Ross gives the following example (Ross, 1967: ex. 4.122):

(3.108) a. Students who fail the final exam or who do not do the reading will be executed.

b.  

```
  S_0
     \   /
    /   /
  VP  S_1
     \  /
    /   /
  VP  S_2
       \  /
       /   /
  NP  VP
   students  fail the final exam
```

He claims that (3.108a) must come from an across-the-board application of Relative Clause Formation to underlying structure (3.108b), rather than from an application of Conjunction Reduction to (3.108c) because (3.108a) and (3.108c) are not synonymous.

(3.108) c. Students who fail the final exam will be executed or students who do not do the reading will be executed.

Under Ross's analysis, (3.90) and (3.98) will act to block (3.108a)! At the end of the $S_1$ cycle, $S_2$-island will be as shown in (3.108b). But at the end of the $S_0$ cycle, $S_1$-island = \{
WH-students (1), fail, \ldots
WH-students (2),
don't, \ldots\} and both $S_2$-island and $S_3$-island overlap with $S_1$-island—-a serious counterexample to my proposals.
I wish to suggest that there is another source of (3.108a) that does not require Relative Clause Formation to operate across-the-board, namely (3.108d).

(3.108) d.

\[
\begin{array}{c}
\text{S}_0 \\
\text{VP} \\
\text{NP}_1 \text{ or } \text{NP}_2 \\
\text{NP} \text{ S}_1 \text{ students fail the final exam} \\
\text{NP} \text{ S}_2 \text{ students do not do the reading} \\
\end{array}
\]

On the \text{S}_0 cycle, Relative Clause Formation applies twice (but not across-the-board) to give (3.108e).

(3.108) e.

\[
\begin{array}{c}
\text{S}_0 \\
\text{VP} \\
\text{NP}_1 \text{ or } \text{NP}_2 \\
\text{NP} \text{ S}_1 \text{ students WH-students ...} \\
\text{NP} \text{ S}_2 \text{ students WH-students ...} \\
\end{array}
\]

Then Conjunction Reduction applies to give (3.108f), which is the same as (3.108a).
(3.108) f.

\[ \text{S}_0 \]

\[ \text{NP}' \]

\[ \text{NP} \]

\[ \text{NP} \]

\[ \cdots \]

\[ \text{students} \]

\[ \text{NP}_1 \]

\[ \text{NP}_2 \]

\[ \text{S}_1 \]

\[ \text{WH-students} \]

\[ \cdots \]

\[ \text{WH-students} \]

The obvious advantage of this analysis is to simplify the grammar by eliminating the need for Relative Clause Formation to operate across-the-board—a strong argument in view of the kind of Structural Description an across-the-board rule needs. (See footnote 28 for Ross's formulation of Conjunction Reduction.)

A second advantage is that (3.90) and (3.98) no longer predict ungrammatical results. At the end of the \[ \text{S}_0 \] cycle, \( S_1 \)-island = \{WH-students, fail, the final exam\}; \( S_2 \)-island = \{WH-students, don't do, the reading\}; \( \text{NP}_0 \)-island = \{\( S_1 \)-island, \( S_2 \)-island\}, \( \text{NP}' \)-island = \{students, \( \text{NP}_0 \)-island\}. There is no overlapping.

As for (3.108b), it underlies sentence (3.108g).

(3.108) g. Students who fail the final exam or do not do the reading will be executed.

It might be claimed that (3.108g) results from an across-the-board application of Relative Clause Formation—giving (3.108a)—followed eventually by an across-the-board
application of WH-REL. These complications to the rules mentioned are unnecessary, especially in view of the necessity of a rule of Conjunction Reduction irrespectively.

Referring to (3.108b), on the $S_1$-cycle Conjunction Reduction applies to give (3.108h):

\[(3.108) \text{h.} \]

\[
\begin{array}{c}
S_0 \\
| \\
NP \\
| \\
NP \quad S' \\
| \\
students \quad \text{NP} \\
| \\
students \quad \text{S}_2 \\
| \\
\text{fail the final exam} \\
| \\
\text{do not do the reading} \\
| \\
\text{will be executed} \\
S_1 \quad S_3
\end{array}
\]

On the $S_0$ cycle, Relative Clause Formation applies, giving, ultimately, (3.108g).

This analysis has the added advantage of interfacing correctly with the island constraints. By the end of the $S_1$ cycle, $S_2$-island = \{fail, the final exam\} and $S_3$-island = \{do not do, the reading\}; $S_1$-island = \{$S_2$-island, $S_3$-island\}; $S'$-island = \{WH-students, $S_1$-island\} and clearly no islands overlap.

Of course the result of "across-the-board" rule applications is still subject to the "usual" island constraint. Thus we expect the grammaticality of (3.108i) to be lowered somewhat--irrespective of the across-the-board rule application.

170
(3.108) i. The girl that I resented that John hit and Bill kicked left.

At the end of the $S_2$ cycle we have:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Thus $S'$-island, $S_2$-island, $S_3$-island and $S_4$-island are as shown. At the end of the $S_0$ cycle, $S_1$-island = \{I, resented, John, hit, Bill, kicked, WH-girl\}. Thus $S'$-island and $S_1$-island overlap, predicting ungrammaticality approximately proportionate to the membership of the girl in $S'$-island.

This membership will be equal to that of an NP-object in an S-complement of resent—as shown in the above diagram—and as we would expect. These are the correct results.

We have shown that the island principles only lower the grammaticality of strings that result from "across-the-board" relativization (which we showed was not across-the-board in the sense that Conjunction Reduction "right" cases. We achieved this result by positing a derivational step of Conjunction Reduction in all cases. In fact, all apparent instances of across-the-board application of some particular rule can be analyzed as involving the operation of Conjunction Reduction.

We illustrate the point further by considering some clefted constructs.

(3.109) a. What Joanne smokes, Robert cultivates and Guinevere fertilizes is pot.

It's pot which Joanne smokes, Robert cultivates and Guinevere fertilizes.
Conjunction Reduction applies on the $S_1$ cycle to produce:

(3.109) b.

On the $S_0$ cycle, the rules of Pseudocleft and (perhaps)
Clefting take place; there will be, clearly, no overlapping islands.

The results discussed in this subsection so far provide
significant support for my hypothesis. It is therefore im-
portant to protect my claim that all across-the-board rule
applications involve the rule Conjunction Reduction.
Grosu (1972: Appendix four) gives three arguments against the stand that Conjunction Reduction is the only across-the-board rule, and that all apparent across-the-board applications of rules such as Relative Clause Formation, Pseudoclefiting, etc., actually involve Conjunction Reduction.

His first argument is that when Conjunction Reduction produces (3.110b) from (3.110a), the pause breaks, marked by the commas, are obligatory, whereas in (3.110c) there are no pause breaks. He assumes, for reasons mysterious, that if Pseudoclefiting operated across-the-board the absence of pauses would be automatically accounted for.

(3.110) a. John likes potatoes and Mary hates potatoes.
    b. John likes, and Mary hates, potatoes.
    c. What John likes and Mary hates are potatoes.

(3.110) b.

```
S
  /\   \                     /\   \   \  
S'  S0  NP                  S0  S1  S2
    /\    \                  /\    /\   \  
   S1    and                S1    S2
   John likes              Mary hates
```

```
S
  /\   \   \                     /\   \   \  
NP  S0  VP                   S0  S1  S2  NP
    /\    \     \                  /\    /\   \  
   what  and  be               S1    S2
   John likes                Mary hates
   potatoes
```

174
The rule that governs pauses, being a phonological rule, operates on the level of surface structure. The rule must be capable of distinguishing (3.110b) from (3.110c) (probably by means of the verb be). But whatever the case, it can't possibly matter whether Conjunction Reduction is an intermediate step, or if Pseudoclefting itself operates across-the-board: the same surface structure will be produced. If Pseudocleft operates across-the-board (3.110c) will be produced directly. If Conjunction Reduction is involved, as in the examples (3.109), then the following intermediate structure is produced:

```
  S
 /\   /
|  |  |
NP S'  VP
    S  \\
      S  \\
    S'  \\
  S   S  \\
   S   S
 P   P
and potatoes
```

When Pseudocleft removes potatoes into the VP, either S' or So will be pruned by the one pruning convention everyone agrees with: \( X \rightarrow \frac{X}{X} \). This will result in the structure shown in (3.110c).

Grosu's second argument is that Conjunction Reduction can only touch initial and final elements of coordinate terms, but rules such as Clefting, Relative Clause Formation, etc. can touch medial elements as well. Thus sentences
like (3.110d) would have no source, since (3.110e) is ungrammatical.

(3.110) d. It's John who I saw in Paris yesterday and
you saw in London the day before yesterday.

e. I saw in Paris yesterday, and you saw in
London the day before yesterday, John.

Grosu may have gotten his idea that Conjunction Reduc-
tion only applies to "initial and final elements of coordi-
nate terms" from Ross's incorrect formulation of the rule in
his thesis (see footnote 28). There are strange constraints
on Conjunction Reduction and one of them has to do with the
phonological heaviness of the element that is reordered.
Thus (3.110f) seems as bad to me as (3.110e).

(3.110) f. Grandma gave several boxes of chocolate
covered maraschino cherries to, and Grandpa
rolled a half dozen good fat numbers for,
John.

Moreover, (3.110g) sounds okay to me.

(3.110) g. I saw in Paris yesterday, and you saw in
London the day before yesterday, the man
that hijacked the Metro to Maubert Mutualité.

The operation of Conjunction Reduction is replete with
mystery. Some kind of output condition will have to lower
the acceptability of sentences like (3.110ef). If a rule
such as Clefting intervenes, however, the string can be
"saved," as (3.110d) is a "save" of (3.110e).

Grosu's third argument is to note that: "across-the-
board chopping is possible, with various degrees of accepta-
bility, in non-coordinate structures. However, Conjunction
Reduction is completely out in such cases..." (p. 221). He
gives as one example the following: (All his "counterexamples" can be handled this way.)

(3.110) g. It's Mary who John likes more than Bill hates.

h. *John likes, more than Bill hates, Mary.

In the next subsection we shall, following Ross, offer the following analysis of sentences such as (3.110g) (we'll be using Relative Clause Formation in that discussion but this doesn't matter).

\[
\begin{align*}
\text{John likes Mary more than Bill hates Mary} & \quad \downarrow \\
& \quad \text{CLEFT, etc.} \\
\text{It's Mary who John likes more than Bill hates Mary} & \quad \downarrow \\
& \quad \text{PRONOMINALIZATION} \\
\text{It's Mary who John likes more than Bill hates her} & \quad \downarrow \\
& \quad \text{PRONOUN DELETION} \\
\text{It's Mary who John likes more than Bill hates.}
\end{align*}
\]

All of Grosu's arguments have been countered and my analysis stands. Thus (3.90) and (3.98), with no additional machinery, handle the across-the-board "exceptions" to the Coordinate Structure Constraint.

3.2.2.2.3 The S that occurs as the "complement" to Comparative sentences has, as we would predict, island-like properties: (3.111b) > (3.111c).

(3.111) a. Guinevere eats popcorn faster than Max eats kibbles.

b. The popcorn which Guinevere eats faster than Max eats kibbles is grown in a hydroponic tank.
c. *The kibbles which Guinevere eats popcorn faster than Max eats were recommended by the Times' food editor.

Ross (1967: 106f) noticed that if the same constituent occurs in both "adjuncts" of a comparative, then it could be removed by an "across-the-board" rule such as Relative Clause Formation—in violation of the Complex NP Constraint—and in violation of (3.98). Thus (3.112) is grammatical; it contains (3.113) as a subpart.

(3.112) The popcorn that Guinevere is devouring faster than Joanne can pop is moving around this kitchen so fast that relativistic effects are becoming noticeable.

(3.113) Guinevere is devouring the popcorn faster than Joanne can pop the popcorn.

We saw in (3.111b) that it is possible to relativize elements out of the main clause of the comparative structure. In view of this fact, then, the grammaticality of (3.114) suggests that it may be an earlier stage in the derivation of (3.112).

(3.114) The popcorn that Guinevere is devouring faster than Joanne can pop it...

To derive (3.114), Relative Clause Formation operates in the ordinary way, relativizing the left-most occurrence of the embedded nominal; later Pronominalization pronominalizes all further rightward occurrences of that same nominal (cf. The man who thought that I wanted to see him). (3.112) is derived not by Relative Clause Formation operating "across-the-board," but by a rule of pronoun deletion whose existence we noted in passing in the last paragraph of 3.2.2.2.2 and which
Ross (1967:108) noted was necessary to derive (3.116) from (3.115):

(3.115) The blintzes which Sasha is gobbling down faster than I can reheat them are extremely tasty.

I suspect that the contract which I wanted to peruse before filing it away may have some loopholes.

The curtain which Fred tore in rolling it up was the kind gift of my maternal Aunt Priscilla.

This rock is too heavy for me to pick it up.

(3.116) The blintzes which Sasha is gobbling down faster than I can reheat $\emptyset$ are extremely tasty.

I suspect that the contract which I wanted to peruse before filing $\emptyset$ away may have some loopholes.

The curtain which Fred tore in rolling $\emptyset$ up was the kind gift of my maternal Aunt Priscilla.

This rock is too heavy for me to pick $\emptyset$ up.

It remains for us to validate (3.98) and (3.90) for the comparative strings we have been discussing. Pretending to know more about the derived constituent structures of comparatives than I really do, let's look at an underlying structure of (3.111bc):

179
(3.117)

\[
\begin{array}{c}
\text{NP} \\
\text{the} \\
\{\text{popcorn (b)}\} \\
\{\text{kibbles (c)}\} \\
\text{S} \\
\text{Guinevere} \\
\text{V} \\
\text{eats} \\
\text{popcorn} \\
\text{faster} \\
\text{S} \\
\text{than} \\
\text{Max} \\
\text{eats kibbles}
\end{array}
\]

At the end of the \(S_1\) cycle, \(S_2\)-island = \{Max, eats, kibbles\}. At the end of the \(S_0\) cycle, if popcorn is relativized as in (3.111b), then \(S_1\)-island = \{Guinevere, eats, WH-popcorn, Max, eats, kibbles\} and \(S_1\)-island contains \(S_2\)-island. If, however, kibbles is relativized as in (3.111c), then \(S_1\)-island = \{Guinevere, ... WH-kibbles\} and the islands overlap. Since \(S_2\) is a relatively weak island, the degree of overlapping is relatively low, and the grammaticality of (3.111c) is relatively high as strings violating the island constraint go.

We can also show that (3.112) does not violate the island constraint:
(3.118)

\[
\begin{array}{c}
S_0 \\
NP \\
the \\
popcorn
\end{array}
\quad 
\begin{array}{c}
VP \\
S_1 \\
N \\
Guinevere
\end{array}
\quad 
\begin{array}{c}
NP \\
is \\
devouring
\end{array}
\quad 
\begin{array}{c}
VP \\
the \\
popcorn
\end{array}
\quad 
\begin{array}{c}
? \\
S_2 \\
NP \\
Joanne
\end{array}
\quad 
\begin{array}{c}
VP \\
than
\end{array}
\quad 
\begin{array}{c}
NP \\
can \\
the
\end{array}
\quad 
\begin{array}{c}
popcorn \\
popcorn
\end{array}
\]

\[S_2\text{-island} = \{\text{Joanne, can, pop, the popcorn}\}; \quad S_1\text{-island} = \{\text{Guinevere, is devouring, WH-popcorn, \ldots Joanne, can, pop, the popcorn}\},\] and we see that \(S_1\text{-island}\) contains \(S_2\text{-island}\), a correct result. Thus the principles (3.98) and (3.90) predict the correct results in all cases.

3.2.2.3.1 Urmson (1963), Hooper (1973) and Hooper and Thompson (1973) have all pointed out and discussed a large class of verbs that allow, on one reading, a parenthetical interpretation. They would claim that (3.119) are ambiguous.

(3.119) \(\begin{cases} \text{We guess} \\ \text{suppose} \\ \text{believe} \end{cases}\) that Ronald sent Spiro a lock of his hair.
In one sense (3.119) indicates a state of supposition or belief on the part of "we," but in another (parenthetical) sense the main verb is reduced semantically to the point where it merely indicates the speaker's attitude to the truth of the proposition expressed by the verb's complement.

The crucial property of parenthetical verbs is that they may end up in various surface positions,\(^{29}\) as illustrated in (3.120):

\((3.120)\) a. Ronald sent Spiro a lock of his hair, we guess.

b. Ronald sent a lock of his hair, we believe, to Spiro.

c. Ronald, we suppose, sent Spiro a lock of his hair.

Note that (3.120c) does not express a supposition; only the parenthetical reading is possible in any of (3.120).

On the other hand, we can suppress the parenthetical reading by applying some island-forming process to an element of (3.119):

\((3.121)\) a. The lock of hair which we believe that Ronald sent Spiro came from the private collection of the Lieutenant Governor.\(^{30}\)

b. It's Spiro whom we suppose that Ronald sent a lock of his hair.

c. It's us who believe that Ronald sent Spiro a lock of his hair.

These data suggest that sentences with parentheticals contain islands. The island constraint would lower the grammaticality of (3.121) relative to the parenthetical reading, allowing the non-parenthetical reading to dominate.
An analysis of the parenthetical and its "complement" into a coordinate-like structure is consistent with the properties of parentheticals discussed by Hooper (1973), and accounts for the island-like behavior we have observed. Thus an underlying structure of the parenthetical reading of (3.120) would be (3.122):

\[
\text{S_0}
\begin{array}{c}
\text{S_1} \\
\text{we} \\
\begin{cases}
\text{guess} \\
\text{suppose} \\
\text{believe}
\end{cases}
\end{array}
\begin{array}{c}
\text{S_2}
\end{array}
\text{returned Spiro a lock of his hair}
\]

The S-islands would be strengthened by virtue of the fact that each is in a different "mode" of communication. The parenthetical is a commentary on the speaker's attitude toward the semantic content of the sentence \(S_2\), and \(S_2\) is the associated assertion.

We should digress slightly to note that the utter badness of (3.123) is probably due to rule ordering rather than a violation of the island constraint:

\[
\text{(3.123) } \text{it's us who Ronald, guess, that sent Spiro a lock of his hair.}
\]

The rule of niching that inserts the parenthetical into its assertive adjunct is a post-cyclic rule that must operate on unmutated parenthetical S-nodes. The fact that niching is a root transformation supports the claim that it is post-cyclic:
(3.124) *I regret that Ronald sent a lock of his hair to Spiro, I believe.

While no verb I know is unambiguously parenthetical, some verbs are more strongly parenthetical than others. **Guess**, **imagine**, and **bet** are three such verbs, accounting for the strangeness of (3.125), where the "literal" reading of the verbs is hard to get, yet forced by Question Formation.

(3.125) Who do you **{guess imagine bet}** is already planning on getting the XXII amendment repealed?

3.2.2.3.2 We carry our discussion of parentheticals over to adverbial clauses. We've already noted, following Emonds (1970), that adverbial subordinate clauses are (at some level) prepositional phrases, and that words such as **because**, **unless**, etc. are prepositions that take an S-complement.

Rutherford (1970) noted that there are two ways in which adverbial clauses may function. One of these he called restrictive (cf. (3.126)), the other non-restrictive (cf. (3.127)).

(3.126) He's not coming to class because he's sick.

He'll take his umbrella in case it rains.

(3.127) He's not coming to class, because he just called from jail.

He'll take his umbrella, in case you're wondering.

In fact, the non-restrictive cases are instances of parentheticals. Structurally, it has been shown by Grosu (1972) and others that ordinary adverbial clauses may be attached to VP-
nodes, but parenthetical adverbial clauses must always be attached to the root-S. Thus ordinary adverbials occur in structures like (3.128); parenthetical adverbials in structures like (3.129):

(3.128)

```
S
  VP
    PP
      +ADVERBIAL
        P
        S
```

(3.129)

```
S
  PP
    +ADVERBIAL
      +PARENTETICAL
        P
        S
```

Our hypotheses would make the following predictions:

(i) a. Island forming processes may not affect elements in adverbial clauses. (Because they form fairly strong islands, as we observed in 2.3.3.1.)

b. Moreover, since parentheticality strengthens islands, violations of island constraints are worse when parentheticals are involved.

(ii) Elements in the main clause of structures of type (3.128) may be operated on by island forming processes.

(iii) But elements in the main clause of structures of
type (3.129) may not be operated on by island forming processes.

Here are some data in support of these predictions:

(3.130a) > (3.130b) and (3.131) > (3.132).

(i) (3.130) a. It's Mary he's got lockjaw because he kissed.

b. It's the Police I'm in jail, for you told.

(ii) (3.131) It's John who's got lockjaw because he kissed Mary.

(3.132) It's me who's in jail, for you told the police.

While it's true that (3.98) and (3.90) will account for all these facts, the constraint mentioned in footnote 31 that parenthetical adverbials must always be attached to the root-S, also account for the facts stated in (iii) and in (i)b. For example, to show how (iii) would be accounted for by our hypotheses, consider (3.133).

(3.133) *The person who won't be coming to class, for Mary just phoned from Katmandu, is John.

```
S_0
   NP
     the N
         person

S_1
     PP
       +ADVERBIAL
         +PARENTHEtical
           P
             for
               Mary just phoned from Katmandu

S_2
     the person won't be coming to class

S_3
```

186

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
At the end of the $S_0$ cycle, $S_1$-island = \{WH-person, ... $S_3$\}, and it overlaps with $S_2$-island, which is as shown in the tree. However, (3.133) is also excluded on the grounds that $S_1$, to which the parenthetical adverbial is attached, is not a root-$S$.

My hypotheses are consistent with fact (ii):

(3.134) The person who won't be coming to class because Mary is sick is John.

At the end of the $S_0$ cycle, $S_1$-island = \{WH-person, ... $S_2$-island\} and there is no overlapping, so (3.134) is predictably grammatical.

To show that the hypotheses account for the predictions made in (i), consider the data in (3.130). Here is a tree diagram of (3.130a):

187
(3.130a)

\[ S_0 \]

\[ \text{NP} \]

\[ S_1 \]

\[ \text{NP} \]

\[ \text{VP} \]

\[ \text{be} \]

\[ \text{NP} \]

\[ \triangle \]

\[ \text{NP} \]

\[ \text{V} \]

\[ \text{NP} \]

\[ \text{PP} \]

\[ +\text{ADVERBIAL} \]

\[ \text{he} \]

\[ \text{has got} \]

\[ \text{lockjaw} \]

\[ \text{P} \]

\[ \text{S}_2 \]

\[ \text{because} \]

\[ \text{he kissed Mary} \]

\[ S_2\text{-island} = \{\text{he, kissed, Mary}\}, \text{ but } S_1\text{-island, which forms after Mary has been promoted into the VP, } = \{\text{he, has got, lockjaw, he, kissed}\} \text{ and the two islands overlap, the strength of overlapping being approximately proportionate to the strength of Mary's membership in } S_2. \]

Now consider (3.130b):

\[ S_0 \]

\[ \text{NP} \]

\[ S_1 \]

\[ \text{S}_2 \]

\[ \text{PP} \]

\[ +\text{ADVERBIAL} \]

\[ +\text{PARENTHETICAL} \]

\[ \text{NP} \]

\[ \text{VP} \]

\[ \text{I} \]

\[ \text{V} \]

\[ \text{PP} \]

\[ \text{P} \]

\[ \text{S}_3 \]

\[ \text{am} \]

\[ \text{in jail for} \]

\[ \text{you told the police} \]

188
$S_1$-island = $\{S_2, \text{you, told}\}$ and $S_3$-island = $\{\text{you, told, the police}\}$ and the two islands overlap. Here the strength of overlapping is greater than in the previous case because the parenthetical $S_3$ forms a stronger island.

While it's certainly true that (3.130a) < (3.130b), in confirmation of the island principles, it is difficult to prove that this difference is to be ascribed to (3.98) because an additional factor adds to the ungrammaticalness of (3.130b), namely the constraint that parenthetical adverbials be attached under the root $S$.

Despite the "noise" of the parenthetical-root-$S$ constraint, we have seen in this section further confirmation of (3.98) and (3.90).

3.2.2.3.3 I have included non-restrictive relative clauses in the section on parentheticals because it is my belief that such clauses stand in the same relation to constituents they modify as parenthetical adverbials to sentences. (And similarly I would claim that restrictive relative clauses are analogous to restrictive adverbial clauses—a claim that certain linguists have made literal by positing underlying restrictive relative clauses for adverbial clauses. I won't pursue the "restrictive" aspect of this interesting question.)

My first argument in support of this hypothesis is that semantically non-restrictives appear to be parenthetical remarks. A non-restrictive relative clause can always be paraphrased by an overtly parenthetical clause:
(3.135) a. The sun, which is ninety-three million miles away, zaps homing pigeons with solar flares.

The sun (and it is ninety-three million miles away), zaps homing pigeons with solar flares.

b. John, who is stronger than a locomotive, brushes his teeth with carpet tacks.

John (and he is stronger than a locomotive) brushes his teeth with carpet tacks.

c. I tried to kiss that girl, who turned out to be Mrs. Nick Dickson.

I tried to kiss that girl (and she turned out to be Mrs. Nick Dickson).

In the second place, NP's that don't allow such overt parentheticals because of various restrictions on the grammar never take non-restrictive relative clauses. 32

(3.136) a. *Anyone (and \{ anyone \} is a human being) cannot he be operated on.

*Anyone, who is a human being, cannot be operated on.

b. \{ Each \} card (and \{ it \} had a picture on it), was dealt face up.

\{ Each \} card, which had a picture on it, was dealt face up.

c. *He didn't write a paper (and it was published in Language).

*He didn't write a paper, which was published in Language.
d. *Did he find a wombat (and \{ \text{the wombat} \} \text{ hides under the sofa?})
   *Did he find a wombat, which hides under the sofa?

e. She is a linguist (and \{ \text{*he studies language} \} \neq
   \{ \text{*it studies she language} \})

*She is a linguist, who studies language.

The third argument is that whatever predicates require that parenthetical adverbials must be attached to a root-S also require that sentence modifying non-restrictives modify the root-S.

(3.137) John is a raving lunatic, for he keeps listening to Bruckner's 9th.

*I regret \{ \text{It is likely} \} that John is a raving lunatic,
for he keeps listening to Bruckner's 9th.

(3.138) John is a raving lunatic, which bothers Mary.

*I regret \{ \text{It is likely} \} that John is a raving lunatic,
which bothers Mary.33

And where parenthetical adverbials needn't be attached to a root-S (e.g., when the main verb has a parenthetical reading, see 3.2.2.3.1), neither must a non-restrictive relative.

(3.139) I suppose John is a raving lunatic, for he keeps listening to Bruckner's 9th.

I suppose John is a raving lunatic, which bothers Mary.

A fourth argument is that just as we noticed that parenthetical adverbials couldn't be reordered by Clefting, or by Adverbial Clause Preposing, although non-parenthetical

191
adverbs could be, non-restrictive relative clauses cannot be reordered by the rule of Extraposition from NP, although restrictive relative clauses can be:

(3.140) Because I love her, I put her vitamins out every morning.

*For I love her, I put her vitamins out every morning.

(3.141) A man walked by and said hello whom I had known in college.

*John walked by and said hello, whom I had known in college.

The claim which I now make, that non-restrictive relative clauses are parentheticals, is well supported by these four arguments and explains why non-restrictives form such strong islands: just as parenthetical adverbials form particularly strong islands due to their parentheticalness, so do non-restrictive relatives. 34

We also note that when a non-restrictive modifies a sentence, that sentence behaves like an island. Similar behavior vis-a-vis parenthetical adverbials was observed in 3.2.3.2. In that case, as in this one, both the island constraint and the "root-S" constraint conspire to lower the grammaticality of the result of applying an island forming process to the main S-clause.

As far as showing details, non-restrictive relative clause formation is structurally similar enough to restrictive clause formation—of which we've seen many examples—to permit us to forego all but one illustration. We choose a case where a whole sentence is modified:

192
(3.142) a. John is insane, which bothers Harry.

b. *It was Harry who John is insane which bothers.

At the end of the $S_1$ cycle, $S_3$-island = \{WH-$S_4$, bothers, Harry\}. At the end of the $S_0$ cycle, $S_1$-island = \{\$2, WH-$S_4$, bothers\} and the islands overlap, predicting the ungrammaticality of (3.142b). Since $S_3$ is a very strong island, (3.142b) is particularly bad.

3.2.2.4.0 We have been illustrating (3.98) and (3.90) in English by means of cleft-sentence structures, relative clause structures, topicalized structures and so on. In this section we shall investigate another island forming process that we have used, namely Constituent Question Formation. We shall look at English, Thai, Korean and Mandarin.

This process is important to us because it suggests one of the crucial theoretical points of this dissertation—that island constraints are constraints on the form of trees.
rather than constraints on rules. Let us first see how things work out in English.

Consider (3.143):

(3.143) What did John find out that Sally was selling?

At the end of the $S_0$ cycle, $S_1$-island = \{Sally, was selling, WH\}. When (3.98) applies at surface structure, $S_0$-island = \{John, found out, Sally, was, selling\}. The islands overlap, but $S_1$-island is quite weak, so (3.143) is not too bad.

To maintain a purely syntactic approach I must suppose that the dcs after WH-Q has taken place is something like (3.144):

(3.144)
That this is so has been argued by Ross, but his sole argument is that such a dcs is needed to make the CNPC work in cases such as *Who and Jill went up the hill? Unfortunately, my sole argument reduces to an analogue of Ross's—I need it to make my island constraint work. (If there was no S' in (3.144), then WH would belong to S_o-island, and we would have containment instead of overlapping.)

It's futile to struggle with argumentation dealing with this question since we intend to consider Constituent Question Formation in languages where a dcs such as (3.144) is out of the question.

Up to this point I have carried out my promise to give constraints and principles that operate on the "visible" syntax—although I noted the influence that semantics had in the form of foreground/background, presupposition, etc. Constituent Questions cannot be handled in this way.

When a speaker formulates a constituent question such as Who ate my porridge? it is accompanied by the presupposition "someone ate my porridge." It is not a simple matter to determine where and how in the grammar the presupposition is to be represented. It may be represented directly as lexical information is by means of feature notation, or it may be derivable, as the notion "subject" or "clause" is. In any case, as I argued in 2.3.1, Constituent Question Formation is a foregrounding operation in which the questioned element(s) is foregrounded while the rest of the sentence is
backgrounded, and it is just the presupposition that is
backgrounded since it, indeed, is "the rest of the sentence."

I am claiming, then, that the presupposition that ac-
companies Constituent Question Formation is an island, and
though it need not be represented overtly in the syntax, it
may overlap with islands that are represented overtly. This
kind of "glitch" is inevitable when a partly semantic notion
such as "island" is treated purely syntactically, as we are
doing for purposes of expostulation.

Let us look at a couple of examples from English before
leaving that language. It is ungrammatical to question ele-
ments of relative clauses because the strong relative clause
island will overlap with the constituent question island:
Consider:

(3.145) *Who does John know a man who hit?

\[
\begin{tikzpicture}
    \node (S0) {S} at (0,0) {};
    \node (NP) at (-2,-1) {NP} child {node (John) {John}};
    \node (VP) at (2,-1) {VP} child {node (V) {V} child {node (knows) {knows}}};
    \node (NP1) at (1,0) {NP} child {node (a) {a}} child {node (N) {N} child {node (S1) {S}} child {node (man) {man}} child {node (NP2) {NP}} child {node (VP2) {VP}} child {node (hit) {hit}} child {node (WH) {WH}}};
    \draw (S0) -- (NP); \draw (S0) -- (VP);
\end{tikzpicture}
\]
$S_1$-island = \{a man, hit, WH\}; CQ-island = \{John, knows, a man, who, hit, someone\}. The overlapping by the lower, strong island $S_1$ predicts that (3.145) will be quite bad.

As a second example we note that we cannot topicalize into a sentence that is subject to Constituent Question Formation.

(3.146) His mother, who sent a letter (to)?

The CQ-island = \{someone, sent, a letter, his mother\}. After topicalization has taken place, $S_o$ will become an island: $S_o$-island = \{WH, sent, a letter\}. The overlapping by the lower CQ-island, which is not a very strong island, results in the somewhat ungrammatical (3.146). Note that (3.145) < (3.146) because relative clause islands are stronger than constituent question islands.

We have been using Constituent Question Formation throughout this work to "test" islands. We noted how constituent questions interact with "squishy" data to give the expected clines of grammaticality. In all these cases we were actually dealing with overlapping islands, and principle (3.98) was at work.
Two examples from Thai will suffice to demonstrate the island constraint. We noted in Chapter two that elements of Thai relative clauses cannot be questioned. The relevant datum (2.28) is reproduced here as (3.147).

(3.147) a. phom chão:b dég thî: phu:cha:j hên
I like boy COMP man see

I like the boy that the man saw.

b. *khun chão:b dég thî: khraj hên
who like boy COMP who see

*Whom do you like the boy that saw.

The relative clause island is \{phu:cha:j, hên, dég\}; the CQ-island is \{khun, chão:b, dég, "someone," hên\}, and the islands overlap.

We also noted in 2.2.3 that it was better in Thai, Korean and Ngizim to question rightmost elements of a series of conjoined NP's than non-rightmost elements. The relevant data were (2.48)-(2.50). We will exemplify this with Thai:

(3.148a) > (3.148b).

(3.148) a. khaw chão:b hua:hm lêe ?araj
he like onion and what

What does he like onions and?

b. khaw chão:b ?araj lêe taengmo:
he like what and watermelon

What does he like and watermelon.

The NP-coordinate island (NP-islands are discussed in 3.3) in (3.148a) is \{hua:hm, ?araj\} and the CQ-island is \{khaw, chão:b, hua:hm, "something"\}. The element ?araj is involved in the overlapping, but its membership coefficient
in the NP-island of \((3.148a)\) -- \(n(?\text{āraj})\) -- is less than its membership coefficient in the NP-island of \((3.148b)\) -- \(m(?\text{āraj})\). In \((3.148b)\) the NP coordinate island is \{?\text{āraj}, taengmo:\} and the CQ-island is \{kh\text{āw}, ch\text{ā}:b, "something," taengmo:\}. Since \(n(?\text{āraj})<m(?\text{āraj})\), we would predict that \((3.148a)> (3.148b)\), which of course it is. These results can be reproduced in Korean, Ngizim and the Ozark dialect of English.

In 2.2.3, the datum \((2.45)\) showed that questioning the subject of a relative clause in Mandarin is worse than questioning an object. Let us see how the details work out:

\[(3.149)\]  
\[
a. \quad \text{tā xīhuan dài shémma de háizi}  
\quad \text{he like wear what rel. part. child}  
\]

What does he like the child who is wearing?

\[
b. \quad \text{tā xīhuan shéi dài de yănjing}  
\quad \text{he like who wear rel part. glasses}  
\]

Who does he like the glasses that is wearing?

In \((3.149a)\) the relative clause island is \{dài, shémma, háizi\}, and in \((3.149b)\) it is \{shéi, dài, yănjing\}. The membership coefficient of shéi in its island -- \(m(\text{shéi})\) -- exceeds the membership coefficient of shémma in its island -- \(n(\text{shémma})\) -- in accordance with the results of section 3.1.3.1.1.1. In \((3.149a)\) the CQ-island = \{tā, xīhuăn, dài, "something," háizi\} and in \((3.149b)\) it is \{tā, xīhuăn, "someone," dài, yănjing\}. The difference in overlapping is just the difference \(m(\text{shéi}) - n(\text{shémma})\), which shows up in the fact that \((3.149a) > (3.149b)\).
We noticed that (3.146), where an object is preposed out of a constituent question structure, was somewhat ungrammatical in English. The mild ungrammaticalness was due to the fact that CQ-islands are fairly weak, and in such structures as (3.146) they are the crucial island in determining overlapping. In Mandarin, sentences analogous to (3.146) are apparently okay, indicating that CQ-islands are quite weak in Mandarin and that (some) speakers have a tolerance for the overlapping islands that we shall now illustrate.

(3.150) shū, nǐ xiǎng shéi xǐhuān kàn
book you think who like read

Books, who do you think likes to read.

The CQ-island = {nǐ, xiǎng, "someone," xǐhuān, kàn, shū}; the island created by the foregrounding rule of object preposing is {nǐ, xiǎng, shéi, xǐhuān, kàn}. The overlapping by the weak CQ-island is based principally on the element shū "books," which as an object has weakest membership in its island, helping to account for the grammaticality of (3.150)36

Korean provides a strong counterexample to my hypotheses that I cannot explain. It is possible, according to one informant, to question elements of relative clauses.

(3.151) a. nà-nán Bìll-ka macna-àn sonyoun-lål
    I Bill met boy
    cohaha-nín-ta
like

I like the boy that Bill met.
b. nò-nín nu-ka macna-ín sonyon-líl
   you who met boy

cohaha-nín-nya
   like

*Whom do you like the boy who met?

There is also class of putative counterexamples. Echo questions in English, Thai and Mandarin can question constituents of relative clauses, coordinate structures, and so on. We will illustrate the matter solely in English though the conclusions we draw about echo questions will be universal.

(3.152) a. You stood off Martha Mitchell and who
   for two hours?

b. These are two fragments of horn concerti that
   who never finished?

c. Martha calls Haldeman and Ehrlichman the
   Katzenjammer Kids because who told her not
   to call them Rosencrantz and Guildenstern?

d. It was beans that who told John I liked?

Echo questions are a unique kind of sentence. They involve automatically repeating material just heard, making certain changes in pronouns to keep everything referentially constant, and substituting a question word(s) in the surface structure position of some element unheard, misunderstood, or about which one wishes to express surprise, shock or indignation (plus intonational changes). Not only do echo-questions run roughshod over principles like the island constraints (mine and everyone else's), but they don't even have respect for the integrity of the word. The following

201
sentences, I believe, are grammatical, though they appear strange in written form.

(3.153) a. Speaker A: Did you know that one million to the twentieth power is called a vigintillion?

Speaker B: (It's called) a what-illion?

b. Speaker A: A poltergeist is always devising tricks to play on my husband and me.

Speaker B: A polter-who?

c. Speaker A: I must hire a lawyer and attempt to gain a replevin.

Speaker B: You want to get a re-what?

Netzron (1972) puts echo questions in a class of "pre-generated structures," i.e. "constructions that, in order to be uttered, presuppose that a sentence with a more or less specific structure has been heard in the previous discourse" (p. 99). He goes on to say that pre-generated structures do not necessarily obey the usual constraints in the language, and that "the copier can take more liberties with the string borrowed than its original emitter" (p. 101). He gives evidence in favor of this view.

Most of his evidence is from Hungarian, but there are a number of instances in English where pre-generated structures behave in an exceptional way.

Consider the negative factive sentence in (3.154).

(3.154) The Katzenjammer Kids didn't regret that they had lied.

A normal uttering of (3.154) carries with it the presupposition that the Katzenjammer Kids had lied. However, if
(3.154) is read as an emphatic denial of another person's previous assertion, the necessity of the presupposition vanishes.

(3.155) The Katzenjammer Kids DIDN'T regret they had lied (how could they REGRET anything when everyone knows that there'll be no whitewash at the Whitehouse).

The pregenerated structure (3.155) is exceptional with respect to the usual facts surrounding the use of factive predicates.

As a second instance of pregenerated structures violating a normally pervasive rule or constraint, consider the following examples.

(3.156) a. Speaker A: I wouldn't be bothered making a dress like that to save $15.

Speaker B: Yeh, but for $50 I would be bothered (making a dress like that).

b. Speaker A: Your thesis advisor isn't fit to eat with the pigs.

Speaker B: Au contraire, my thesis advisor is eminently fit to eat with the pigs, so there!

"Negative polarity" items such as "be bothered," "be fit," "could care less," etc. normally cannot occur in a non-negative mode—but as pregenerated structures they can.

As a third example we note that "crossing restrictions" can be violated by pregenerated structures. Normally the string John was shaved by himself is excluded by the grammar, but in reply to the question Who was John shaved by? the answer John was shaved by HIMSELF, with contrastive stress on the reflexive is acceptable.

203
As a fourth example I offer an excerpt from a Dial Soap commercial of May, 1973:

Speaker A: Have you felt the gentle difference?
Speaker B: ?The what difference?

Only in pregenerated structures can adjectives be questioned directly by what (to the extent they can be questioned at all).

As pregenerated structures echo-questions belong to a very narrow, well-defined class of exceptions to the island constraint and therefore do not provide damaging counter-evidence. What the underlying cause is of the idiosyncratic behavior of pregenerated structures is unfortunately not clear to me. 37

3.2.2.5 I have confined the discussion in this work to islands in which NP's overlap. However verbal elements and even S's are subject to movement rules that can result in overlapping islands—and as (3.98) predicts, lowered grammaticality results.

There are two syntactic phenomena in English that reorder sentences in such a way as to have the potential of producing overlapping islands. One of these is sentential relativization, illustrated in (3.157). 38

(3.157) The president stated that there would be no whitewash at the Whitehouse, which no one of sound mind could possibly believe.
At the end of the $S_0$ cycle, $S_2$-island = \{NP$_2$, V$_2$, WH-S$_3$\}, and it is clear that no islands can overlap. (By (3.105) $S_1$ and $S_2$ are disjoint.) Consider, on the other hand, (3.158).

(3.158) a. Jo is mad, which few people realize.
   
   b. Jo is mad, which Jan believes that few people realize.
   
   c. Jo is mad, which Jan believes the claim that few people realize.

Clearly (3.158a) $\succ$ (3.158b) $\succ$ (3.158c); let's look at (b):

205
At the end of the $S_2$ cycle, $S_3$-island = \{few people, realize, $S_4$\}; at the end of the $S_0$ cycle, $S_2$-island = \{Jan, believes, few people, realize, WH-$S_4$\}, and by convention (3.91) of 3.2.1.1 the islands overlap. The strength of overlapping is measured by the strength of $S_4$ in its island $S_3$—which is weak, so (3.158b) is not too bad.

Here is (3.158c):

In this case $S_3$-island is much stronger, and the degree of overlapping (which involves essentially the same elements) is consequently greater, hence (3.158c) is fairly ungrammatical.

The second $S$-reordering process is clefting.

(3.159) a. It's that Jo is mad that few people realize.

or

What few people realize is that Jo is mad.
No islands overlap in this case, but consider:

(3.159) b. It's that Jo is mad that Jan believes the claim that few people realize.

or

What Jan believes the claim that few people realize is that Jo is mad.
At the end of the $S_1$-cycle, $S_2$-island = \{ few people, realize, $S_3$ \}, and at the end of the $S_0$ cycle, $S_1$-island = \{ ... few people, realize \} and there is overlapping, as the low grammaticality of (3.160) leads us to suspect.

Verbal elements can also be involved in overlapping. One instance is the preposing of predicate adjectives, exemplified in (3.160):

(3.160) a. Smart though I believe the claim that my students are, they still think predicate raising is something you do with a straight flush.

It's not clear to me what underlies (3.160): perhaps something like (3.160b) at one stage:

(3.160) b. Though I believe the claim that my students are smart, my students still think...

$S_2$-island will contain smart, but $S_1$-island will not. Those overlapping islands are sensed in the low grammaticality of (3.160a).
Two other cases are parallel in all detail to the two cases involving sententials, as the data in (3.161) suggest.

(3.161) a. If he \{goes \{kicks Mary\}, which I don't think he will, ...

b. What I believed that John decided to try was to write a book.

c. What I believed the claim that John decided to try was to write a book.

It is clear that the island constraint should lower the grammaticality of (3.161c):

```
S₀
  NP
  VP
    S₁
      NP
      VP
        I
        V
        believed
        the
        claim
        S₂
          NP
          VP
            John
            V
            decided
            V
            to try to write
            NP
            a book
```

At the end of the S₁ cycle, S₂-island = \{John, ... write, a book\}, but at the end of the S₀ cycle, S₁-island = \{I, believed, ... John, decided\}. The islands overlap and (3.98) lowers the grammaticality of (3.161c).
3.2.2.6 We have discussed deletion rules to a small extent--English Relative Clause Formation is sometimes a deletion rule, and Relative Clause Formation in Thai and other languages we have investigated is always a deletion rule.

Deletion rules may create overlapping islands, and when they do we expect (3.98) to lower the grammaticality of the results.

Comparative deletion is a typical deletion rule.

(3.162) a. John is taller than Bill *(is).

b. ?John is taller than I believe that Bill*(is).

c. ??John is taller than I believe the claim that Bill *(is).

We expect (3.98) to tell us that (3.162a) > (3.162b) > (3.162c), with (3.162c) pretty bad.

Because comparative deletion is (most likely) a cyclic rule (it must precede Pseudocleft to account for The one who John is taller than is Bill), there will be no overlapping islands in the surface structure of (3.162a):

```
(3.162a)
S₀
  /\  \\
 NP  VP
  |    |
  John V
     |    |
     tall than
```

In (3.162b) there is a mild violation if tall is deleted, and a strong violation if both be and tall are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
deleted. By supposing that be has strong membership in its island we account for why its absence in (3.162b) results in low grammaticality, but its absence in (3.162a) doesn't affect grammaticality at all—a fact that seems otherwise perplexing.

(3.162b)

At the end of the $S_1$ cycle, $S_2$-island = \{Bill, be, tall\}. But at the end of the $S_0$ cycle, $S_1$-island = \{I, believe, Bill, (be)\} and the islands overlap.

In (3.162c), because tall and be are in a stronger N-S island, the overlapping results in a greater lowering of grammaticality. The details are nearly identical to those in (3.162b).

Another deletion rule is the one that relates

(3.163) a. This dog is too stupid for me to play backgammon with this dog.

b. This dog is too stupid for me to play backgammon with.

The rule optionally deletes the identical this dog; if it doesn't apply, Pronominalization produces (3.163c):
c. This dog is too stupid for me to play backgammon with her.

The deletion can produce overlapping islands, and these are of lowered grammaticality, as (3.164) illustrates:

(3.164) This dog is too stupid for me to believe the claim that your cat lost to ?*(her) at backgammon.

\[ S_0 \]
- NP
  - this dog
  - too stupid

\[ S_1 \]
- NP
  - I
- V
  - believe
  - the claim

\[ S_2 \]
- NP
  - your cat
- VP
  - lost to this dog at backgammon

At the end of the \( S_1 \) cycle, \( S_2 \)-island contains this dog, but at the end of the \( S_0 \) cycle, \( S_1 \)-island will not contain that element, so the islands overlap.

There are deletion rules that produce overlapping islands that are not ungrammatical—counterexamples to my hypotheses.

The rule Genitive Head Deletion is one of these; it produces the grammatical (3.165):

(3.165) Haldeman burgled Ehrlichman's files and Ehrlichman hired the man who burgled Haldeman's.
At the end of the $S_2$ cycle, $S_3$-island contains the element *files*, but at the end of the $S_0$ cycle, $S_2$-island does not contain that element and the islands overlap. The exception is narrowly defined, but difficult to account for in a principled way.

To the extent that the following analysis is valid, this counterexample can be avoided: Pronominalization, a post-cyclic rule, converts the second $N$ to *ones*. It affects both $S_2$-island and $S_3$-island equally so $S_2$-island continues to contain $S_3$-island. A deletion rule recognizes that the string *NP's ones* is not allowed and deletes the *ones*, but this post-cyclic deletion doesn't affect the status of the islands. This is the kind of analysis suggested in Postal (1966).
There are two other exceptions that are unquestionably related to each other: these involve VP-deletion and Sluicing.

(3.166) The CIA can violate the Constitution, but the President denied the claim that the FBI can. (VP-deletion)

(3.167) Someone burgled Ellsberg's psychiatrist, and the President knows the man that knows who. (Sluicing)

We will illustrate the overlapping islands with (3.167).

\[
\begin{align*}
S_4 \text{-island contains } & \text{burgled Ellsberg's psychiatrist, but } S_2 \text{-island will not, if Sluicing is not post-cyclic.}
\end{align*}
\]
There is evidence that Sluicing is in fact post-cyclic. It must follow WH-Q, as Ross argues extensively in "Guess Who" (1969). VP-deletion is also post-cyclic, as suggested by (3.168):

\[(3.168) \text{ Guinevere knows } \{ \text{why she should} \} \text{ knock over small girls on roller skates, but I don't remember } \{ \text{why I should} \} \{ \text{how to} \}.\]

VP-deletion applies after WH-Q preposes the adverbial question word how/why to the front of the embedded Q-complement.

With these two rules being post-cyclic, convention (3.90) determines islands in such a way that sentences such as (3.166) and (3.167) are not predicted as ungrammatical, a correct result, and further evidence in favor of the post-cyclicity of the two rules.

I cannot, alas, have it "both ways" and it is important to note that certain sentences that are transformed by Sluicing provide counterevidence to (3.90).

The sentences in (3.169) form a three-way cline of grammaticality:

\[(3.169) a. \text{ She kissed a man who bit one of my friends, } \text{ but Tom doesn't realize which one of my friends she kissed a man who bit.}\]

\[b. \text{ She kissed a man who bit one of my friends, but Tom doesn't realize which one.}\]

\[c. \text{ She gave a nickel to one of my friends, but Tom doesn't realize which one.}\]

(3.169a) is excluded from the language by (3.98); and (3.169c) is included. As things stand, however, (3.169b) is given the

215
same status by (3.98) as (3.169a), for just as Sluicing doesn't affect the islands in (3.167)—a correct result—it fails to affect the islands in (3.169b)—an incorrect result.

If Sluicing is everything Ross (1969) says it is, I can see no way around the counterexample. However the semantic reading of sentences in which an overlapping island has been "sluiced away" is not what Ross claims. Ross would say that the reading of (3.169b) is the same as the "intended" reading of (3.169a). I find that (3.169b) is interpretable as:

(3.170) "She kissed a man who bit one of my friends, but Tom doesn't realize which one of my friends a man bit."

Similarly, while Ross claims the meaning of (3.171a) is (3.171b), I find (3.171c) a more likely interpretation.

(3.171) a. Irv and someone were dancing together, but I don't know who.

b. "Irv and someone were dancing together, but I don't know who Irv and were dancing together.

c. "Irv and someone were dancing together, but I don't know who Irv was dancing with.

The problems of deriving sentences such as (3.171a) or (3.169b) from (3.171c) or (3.170) could be enormous. Even though Ross points out that Sluicing is "sloppy," one hesitates to try and formulate a rule that is sloppy enough to derive (3.171a) or (3.169b). Were the rule to be extended somehow, the extent of sloppiness may account for the somewhat lowered acceptability of the sluiced sentences.
Sluicing is an "identity of sense" deletion rule. A comprehensive understanding of "identity of sense" in terms of our present models of transformational grammar does not exist. As a result, there are many questions unanswered as to the exact scope and operation of this rule. Hopefully a more precise understanding will find the process of Sluicing less in conflict with principle (3.90).

3.2.2.7 The Sentential Subject Constraint (Ross, 1967:134), to the extent that it is not encompassed by the Internal Island Sentential Noun Phrase Constraint (IISNPC), falls within the realm of the island constraint I am proposing, and is therefore superfluous.

The IISNPC was given by Ross (1967:33):

(3.172) Grammatical sentences containing an internal NP which exhaustively dominates S are unacceptable.

It was formulated to account for cases like those in (3.173):

(3.173) a. *Did that John showed up please you?
   b. *That that John showed up pleased her was obvious.
   c. *I want that Bill left to remain a secret.
   d. *I went out with a girl who that John showed up pleased.

As stated, one can immediately find counterexamples to (3.172), as Ross did later on in his thesis, as Grosu (1972:156) did, and as I did (Rodman, 1971:146). For example:

(3.174) a. Bill said that for her to enlist would be impossible.
   b. Is what you ate known to your doctor?
c. Does whoever steals your purse steal trash?

d. Carol was assured that Bob loves Ted by Alice.

e. I explained that Guinevere had merely been hungry to the parents whose twin baby girls that nutty dog had just devoured.

f. Did John's showing up please you?

Later, (3.172) was amended to explicitly exclude gerunds, taking care of (3.174f), but the other cases remained until Ross discovered that IISNPC was actually a fuzzy constraint sensitive to the degree of nouniness of "headless sentence-internal nominal complements" (cf. section 3.1.3.1.4). This fact is illustrated by (3.175) where $a \leq b \leq c \leq d \leq e \leq f \leq g$.

(3.175)

a. Is that he'll stop at every toilet along the way likely?

b. Is that he'll stop at every toilet along the way ridiculous?

c. Is for him to stop at every toilet along the way likely?

d. Is where you live known to the FBI?

e. Does his stopping at every toilet along the way bother you?

f. Will his resisting of the police disturb the neighbors?

g. Will his resistance of the police get him into trouble?

The constraint is considerably more effective on that-S complements than any other kind, becoming gradually less effective from for-to structures on down. This accounts for the counterexamples (3.174abcf), but still fails to explain the high grammaticality of (3.174de). To my knowledge no

218
one has been able to explain these two counterexamples to the IISNPC, each of which represents a fairly large class of counterexamples:

\[
\begin{align*}
\text{(3.176) Carol was} & \begin{cases} \text{assured} \\ \text{advised} \\ \text{convinced} \\ \text{informed} \\ \text{reminded} \\ \text{told} \\ \text{warned} \end{cases} \text{that Bob loves Ted by Alice. (Cf. (3.174d))} \\
\text{(3.177) We} & \begin{cases} \text{confessed} \\ \text{explained} \\ \text{confided} \\ \text{declared} \\ \text{emphasized} \\ \text{indicated} \\ \text{reported} \\ \text{stated} \\ \text{specified} \\ \text{whispered} \end{cases} \text{that we were hungry to the chief. (Cf. (3.174e))} 
\end{align*}
\]

Despite these counterexamples a fuzzy constraint very much like IISNPC is needed in a grammar of English and we will assume its existence.

Ross notes in his Dissertation that it appears that "the SSC ... is a special case of [IISNPC]" but goes on to note "two arguments...[which are]...only to be accountable for if [the SSC] is assumed to be operative in the grammar of English."

The first of these arguments is that (3.178a) is worse than (3.178b). In (3.178a) Ross claims, both the SSC and the IISNPC are violated, but in (3.178b) only IISNPC is violated.

\text{(3.178) a. The hat which that I brought seemed strange to the nurse was a fedora.}
b. The nurse who that I brought this hat seemed strange to was as dumb as a post.

The actual effect of the SSC, to the extent that it is isolatable, is rather mild; these sentences seem approximately equally bad.

The second argument involves the difference in grammaticality between (3.179a) and (3.179b).

(3.179) a. That piano, which the boy's loud playing of drove everyone crazy, was badly out of tune.

b. That piano, which the boy's playing loudly drove everyone crazy, was badly out of tune.

The (b)-sentence is claimed by Ross to be worse than the (a)-sentence. Schematically (3.179a) looks like this:

```
  S
   \--- action nominal
    \-- The boy's loud playing of the piano
       \-- VP
        \-- drove everyone crazy
```

And (3.179b) like this:

```
  S
   \-- NP's Ving
    \-- The boy's playing the piano loudly.
        \-- VP
         \-- drove everyone crazy
```

Ross claims that the action nominal is dominated by an NP node, but that the NP's Ving is dominated by an NP structure. The following tree diagrams are Ross's:
(3.179a)

S
  /   \
 VP  NP
   /   \
  NP   NP
   /   \
  NP   NP
   /   \
  NP  NP
   /   \
  NP  "of the piano"
   /   
  "the boy's loud"
   /   
  N   V
   /   
  "playing"

(3.179b)

S
  /   \
 NP  VP
   /   \
  NP   NP
   /   \
  NP  "playing the piano loudly"

He then explains that part of the reason the (b)-sentence is worse is that the SSC is violated, but in (3.179a), where an S-node is lacking, this would not be the case.

To this last argument, Ross, as well as myself, would now answer that the nodes dominating the structures in question are fuzzy nodes, hovering somewhere between full sentenceness and full nouniness. A principle set out in 3.1.3.1.4 states that nounier structures form stronger islands. Since action nominals are nounier than gerunds,
that principle in isolation predicts that (3.179a) will be worse than (3.179b)! On the other hand, the IISNPC constrains the more sentencelike gerund to a greater extent than it constrains the action nominal. Perhaps other, unknown factors are present. On top of all this, I find it very difficult to determine relative grammaticality between (3.179ab)--not too surprising in view of the opposing forces at work on these strings.

It does not seem to me that this fragile case is sufficient evidence for a constraint as broad and as powerful as the Sentential Subject Constraint--especially in light of what I now have to say.

Returning to the first example, we must note that the sentential subject is an island, and may be subject to certain strengthening. For instance Hooper (1973) notes: "It is true, however, that all sentential complements in subject position are de-emphasized or backgrounded..." She also reminds us that the Kiparsky's noted that sentential subjects are often presupposed. Thus the test of negation shows that the sentential subject is presupposed in (3.180), but the same sentence is not presupposed in object position (cf. (3.181)).

(3.180) a. That Smith had arrived was reported by the UPI.
   b. That Smith had arrived wasn't reported by the UPI.

(3.181) a. The UPI reported that Smith had arrived.
   b. The UPI didn't report that Smith had arrived.
Both these factors strengthen the sentential subject island.

We may conclude, then, that the Sentential Subject Constraint does not exist. Sentential subjects are, however, moderately strengthened islands, and that fact, in conjunction with IISNPC will account for the observed data.

Ross noted that there is no SSC in Japanese, and, in fact, Grosu (1972) notes that it fails to hold in Turkish and Hindi. Grosu suggests that only SVO languages will have a Sentential Subject Constraint (Japanese, Turkish and Hindi are SOV). I am suggesting that no language has a Sentential Subject Constraint. The question is why do SVO languages behave as if there was a Sentential Subject Constraint. I believe I have answered this question for English. Other languages will have to be studied as to language specific details.

3.2.2.8 One of the basic tenets of this work, that all non-root S's are fuzzy islands, is far from being an original idea. In his thesis Ross credits Janet Dean with having suggested the following hypothesis:

(3.182) No element of a subordinate clause may be moved out of that clause.

Ross rejects Dean's proposal for three reasons, which we are obliged to answer.

His first reason is that the constraint would not be strong enough to account for the differences in grammaticality in (3.183) (Ross's (4.251):133), and therefore (3.182) would have to be supplemented by the Sentential Subject Constraint.
(3.183) a. The teacher who the reporters expected that the principal would fire is a crusty old battleax.

b. *The teacher who that the principal would fire was expected by the reporters is a crusty old battleax.

c. The teacher who it was expected by the reporters that the principal would fire is a crusty old battleax.

This situation was dealt with in the last section (3.2.2.7) and is not problematic.

Quoting Ross (1967:139): "The second [reason] is that [(3.182)] would incorrectly exclude all the sentences of (2.23), [given as (3.184) below] which differ among themselves in acceptability, but some of which seem perfectly normal to me."

(3.184) a. He told me about a book which I can't figure out {whether to buy or not,
how to read.
where to obtain.
what to do about.}

b. He told me about a book which I can't figure out {why he read.
?whether I should read.
??when I should read.}

c. Which books did he tell you {whether} he wanted to read.

In 3.1.3.1.3 and 3.1.3.2 we discussed in detail the kinds of sentences given in (3.184), and proposed hypotheses to account for them that are perfectly compatible with the other hypotheses of this Dissertation.

Ross's third reason "is that elements of clauses with POSS-ing or for-to complementizers can be relativized ..."
He gives the following sentence as an illustration.

(3.185) The only hat which it bothers me for her to wear is that old fedora.

Based on our work so far, it might seem that (3.185) should be worse than it is, since for-to complements are nounier than that-S complements, and nouniness is island-strengthening.

What we must also recall is that for-to complements are less clausy than that-S complements, as discussed in footnote 17 in a slightly different context. In the clausiness squish of 3.1.3.2.1, for-to complements come between tenseless that-S clauses and subjected non-finites, and are therefore fairly low in clausiness.

As a first argument in support of this claim, there are some Verb-Verb restrictions—not as strong as with the subjected non-finites, but stronger than with tenseless that-S clauses: most for-to complements must not be stative:

(3.186) *I {argued, pleaded, hoped} for him {to own a cat, to be tall.}

but: It bothers me for my dates to be tall.

Secondly, Equi-NP-Deletion applies obligatorily in most cases of for-to complementization:

(3.187) I am {eager, anxious, willing} {for him} to go.

I {asked, intend, prefer, begged} {for him, *for me} to go.
In the third place, some for-to complements require the "unlike subject constraint."

\[(3.188) \quad \begin{cases}
\text{screamed} \\
\text{prayed} \\
?\emptyset
\end{cases} \quad \{ \text{for him} \} \quad \text{to go.}\]

A fourth argument that for-to complements are low in clausiveness is that semantically they are almost always unfulfilled actions; such semantic restrictions are indications of reduced clausiveness.

\[(3.189) \quad \begin{cases}
\text{pleaded} \\
\text{arranged} \\
\text{yearned}
\end{cases} \quad \text{for him to have already gone.}\]

We conclude that for-to clauses are low in clausiveness. Thus a strengthening factor and a weakening factor converge on for-to complements. The result is a fairly weak island, as Ross noted.

Ross gives no examples of relativization of elements of clauses with a POSS-ing complementizer. Presumably he had something like this in mind:

\[(3.190) \quad a. \quad \text{The girl who I dislike his tickling went ape when I showed her a nude photo of J. Edgar Hoover.}\]

\[b. \quad \text{Who do you approve of my seeing?}\]

I find these "nony" complements to be somewhat resistant to relativization (and other processes that might create overlapping islands)—(3.190) is not entirely grammatical. As with for-to complements, however, I think (3.190) is perhaps better than expected, and for exactly the same reason—gerundive complements are low in clausiveness—occurring in the clausiveness squish close to tenseless that-S clauses.
There are no Verb-Verb restrictions that I am aware of. Equi-NP-Deletion applies with frequency, often obligatorily.

\[(3.191) \quad \text{I\{hated\} regret\{\text{his} \} \text{enjoyed\{\?\text{my} \}} \text{approved\{of\}} \text{leaving early.} \]

And the unlike subject constraint holds for a number of cases:

\[(3.192) \quad \text{I\{predicted\} prevented\{\text{*my} \} \text{awaited\{\text{his} \}} \text{leaving.} \]

The same conclusions apply to POSS-ing complements that applied to for-to complements. That is, an "anti-conspiracy" keeps the islands fairly weak.

The third argument of Ross, then, is accounted for by the proposals made in this chapter and we persist in keeping the hypothesis that all S's are fuzzy islands.

3.2.2.9 As I noted in Rodman (1973), one of the puzzling exceptions to all island constraints I know of, including my own, is the fact that in Japanese it is possible—at least in some instances—to relativize an element of a relative clause. But Japanese is not a pronoun-leaving language!

In a work completely independent of considerations in this Dissertation, Perlmutter has proposed a potential solution to the Japanese counterexample. But we must digress to see precisely what the problem is.

Consider the Japanese sentence in (3.192).

\[(3.192) \quad \text{child petting-was dog died perf} \]

The dog that the child was petting died.

227
Japanese relative clause formation is apparently a deletion operation, as in Thai.

But consider (3.193), in which "child" is relativized out of (3.192):

(3.193) kawaigatte-ita inu ga sine simatta kodomo
petting was dog died perf child

The child that the dog that he was petting died...

(both deleted by R.C. Formation)
At the end of the $S_1$ cycle, $S_2$-island = \{kodomo, kawaigatte-ita\}. At the end of the $S_0$ cycle, ($S_0$ is not shown), $S_1$-island = \{kawaigatte-ita, inu, sine simatta\} and the islands overlap. The grammaticality of (3.193), then, represents the type of counterexample we have to cope with in Japanese.

Perlmutter (1972) has argued that relativization takes place in two stages: the relativized nominal is reduced to a "shadow pronoun" by the rule of Relative Clause Formation, and a later rule of "Shadow Deletion" deletes the shadow pronoun in non-pronoun-leaving languages (I am simplifying the matter down to the relevant points).

Perlmutter argues that Japanese does not have a rule of Shadow Deletion (that is, Japanese is a pronoun-leaving language), but does have a general rule of "Pronoun Drop," which accounts for the surface form of Japanese relative clauses. Moreover, Perlmutter notes that "It is possible to obtain the correct result within transformational theory by ordering Shadow Deletion before Pronoun Drop, or by making Shadow Deletion cyclical and Pronoun Drop post-cyclical." (My underscoring.)

In light of Perlmutter's analysis, let's look at (3.193) again: At the end of the $S_1$ cycle, $S_2$-island = \{kodomo, pro-inu, kawaigatte-ita\}. At the end of the $S_0$ cycle, $S_1$-island = \{pro-kodomo, pro-inu, kawaigatte-ita, inu, sine simatta\}. Like Sluicing and VP-Deletion in English, the post-cyclic rule of Pronoun Drop does not affect the status of the islands, which do not overlap.
To the extent, then, that Perlmutter's analysis is valid, we are able to avoid the counterexample from Japanese.

3.3

3.3.0 In this section we will discuss that part of statement (3.1) relevant to NP's.

(3.1) All non-root S's and NP's are (fuzzy) islands.

We will confine discussion for the most part to English since my experience with NP's in other languages is limited. I don't know if the fact that NP's are fuzzy islands is a universal property of language. I will attempt to establish this fact for English alone.

The principles formulated in sections 3.1 and 3.2 apply equally well to NP's, except of course for the ones that are S-specific. That this is so provides strong evidence in favor of these principles. In particular, the principles (3.90) and (3.98) appear to hold when NP's are treated.

That NP's are indeed (fuzzy) islands is suggested by the following sampling of data, none of which, I believe, are fully grammatical.

(3.194) The accident which the police discovered Mary's body at the place of should have been avoidable.

The furniture which we discussed his interest in went up in price.

Who did John return to the city in spite of the presence of?

The man whom I enjoyed that performance by was Glenn Gould.

The car which the damage to was excessive will have to be junked.
The gun that I saw a man with was loaded.

The professor who the army edited the analysis of the report by was indignant.

Whose did you find book?

That the man left bothers me that I had just recently met.

What did John eat bacon and?

The reason that my proposal that NP's are islands has not, to my knowledge been made before is that there are data such as (3.195) which suggest that such a generalization is spurious.

(3.195) Reports which the government prescribes the height of the lettering on the covers of are invariably boring.

As with S's, NP's are basically weak islands subject to strengthening factors. Rightmost elements of some kinds of NP's will have nearly zero membership in their NP island. These cases are exceptional. All in all, however, (3.1) provides a measure of simplicity while allowing us to account for a large body of data heretofore treated inadequately or not at all.

3.3.1 We first discuss the problems that led Ross (1967) to formulate the Pied Piping convention and the Left Branch Condition. Ross noted that under certain circumstances a constituent which dominated a constituent "mentioned" in a chopping rule can be reordered in the place of the mentioned constituent. Thus WH–REL mentions NP–WH, and accounts directly for (3.196).
(3.196) Charlie discussed a claim which I discovered evidence for \( \leftarrow \) Charlie discussed a claim \( S[I \text{ discovered evidence for WH-claim}]g. \)

But there is a stylistic variant of (3.196) in which the entire NP evidence for WH-claim is preposed.

(3.197) Charlie discussed a claim evidence for which I discovered.

Ross (1967) argues at length that no modification of the rules will accommodate the facts since the process involved is unbounded:

(3.198) \[
\begin{align*}
&\{ \text{Reports which} \\
&\quad \{ \text{Reports the covers of which} \\
&\quad \{ \text{Reports the lettering on the covers of which} \\
&\quad \{ \text{Reports the height of the lettering on the covers of which} \\
&\quad \{ \text{the government prescribes} \\
&\quad \{ \text{the height of the lettering on the covers of} \\
&\quad \{ \text{the height of the lettering on} \\
&\quad \{ \text{the height of} \\
&\quad \{ \emptyset \\
&\text{are invariably boring.}
\end{align*}
\]

His basic insight is that the alternative movements are to be accounted for not by a modification of WH-REL, but by the implementation of a convention on chopping rules--the so-called Pied Piping Convention.

(3.199) **THE PIED PIPING CONVENTION:** Any transformation which is stated in such a way as to effect the reordering of some specified node NP, where this node is preceded and followed by variables in the structural index of the rule, may apply to this NP or to any non-coordinate NP which dominates it, as long as there are no occurrences of any coordinate node, nor of the node S, on the branch connecting the higher node and the specified node.

As one can gather merely by reading the convention, there is a problem in formulating a comprehensive statement.
of the conditions under which Pied Piping operates. Indeed, the facts are so complex that Pied Piping by itself should be considered a major goal of future research.

One complication is that the conditions stated in (3.199) are not entirely accurate, as Postal (1972) points out. (We will discuss one aspect of this when we consider fuzzy nodes in Appendix ) Another problem is that Pied Piping is somewhat dependent on the rule involved although its chief virtue is being "rule independent." Thus Postal (1971:99) has noted that:

Pied Piping is blocked for WH-Q movement when that NP in the proper analysis which is marked (+WH) is the rightmost constituent of some higher NP (which does not directly dominate a preposition).

The data in (3.200) and (3.201) suggest the verity of the above statement.

(3.200) a. Charlie discussed a claim which I discovered evidence for.
   b. Charlie discussed a claim for which I discovered evidence.
   c. Charlie discussed a claim evidence for which I discovered.

(3.201) a. What claim did Charlie discuss evidence for?
   b. For what claim did Charlie discuss evidence?
   c. *Evidence for what claim did Charlie discover?

It also appears that Pied Piping is structure dependent. For most speakers no element may be pied piped in a Q-complement. Right-branching NP's in restrictive relatives are harder to pied pipe than they are in non-restrictives:
(3.202) *I wonder by whom he received an offer.
       *I wonder an offer by whom he received.

(3.203) Reports the covers of which the government
       prescribes the cost...
       ?Any reports the covers of which the government
       prescribes the cost...

One further difficulty is whether prepositions are
preposed by pied piping or by being mentioned optionally in
such rules as WH-Q and WH-REL. For example, the SD of WH-REL
could be formulated as (3.204) or (3.205).

(3.204) \[ W - NP[NP - S[X - NP - Y]_S]_{NP - Z} \]
(3.205) \[ W - NP[NP - S[X - (P) - NP - Y]_S]_{NP - Z} \]

In (3.204), it is presumed that "PP's" are NP's. Complica-
tions such as the fact that sometimes a "PP" is reluctant to
pied pipe (e.g. ?The race of which we demanded an inquiry)
would be a further "glitch" in the Pied Piping Convention.
In (3.205) "PP's" can remain PP's, but the rule would need
certain conditions to account for idiosyncratic cases.

Which of these choices is made affects us here because
if PP's are NP's at the point when islands are defined we
will have to account for why prepositions are productively
stranded—in many cases obligatorily—without a discernable
lowering of grammaticality. If, however, PP's are PP's, then
because only NP's are islands the island constraint will not
act to lower grammaticality whenever the PP overlaps, as
when its preposition is stranded.
Many more difficulties with Pied Piping are discussed in Ross (1967), Postal (1971) and Rođman (1972b).

Apart from the unsteady status of PP's we can divorce ourselves from problems involving the details of Pied Piping and concentrate on the subject at hand. Postal (1971) has distinguished between "two really quite distinct aspects of Ross's proposal:

(3.206) Pied Piping Function
It is a functional property of some transformational rules that they operate on the proper analyses of phrase markers in such a way as to reorder not an NPa mentioned in some term T of the proper analysis but rather some NP in the phrase marker which dominates NPa.

(3.207) Pied Piping Scope
Those rules which manifest this property are characterizable by the condition X and operate in the class of contexts Y.

For our purposes we are interested in (3.206) and how the fact stated there might interact with the island-relevant principles. While (3.207) is a subject of fundamental importance, it involves matters that are too highly idiosyncratic for us to deal with in detail (but see Appendix ).

We noted in 3.1.3.1.1 that nodes within S-islands have stronger membership in the island if they are more to the left. This fact, which I do not fully understand, is perpetuated even more vigorously in NP-islands:

(3.208) Non-right-branching elements in NP's have very strong island membership.

Consider the following datum:

(3.209) *The man whose you saved life slept.
At the end of the $S_1$ cycle, $NP_1$-island = \{man, life\}; at the end of the $S_0$ cycle, $S_1$-island = \{you, save, WH-man, life\} and the two islands overlap. The strong island membership of man is responsible for the high degree of ungrammaticality. Principle (3.208) also accounts for why the head of an N-S construct cannot be relativized, etc.:

(3.210) *The claim which we debated that W.C.Fields was an oblate spheroid can never be proven for sure.\textsuperscript{42}

(3.208) may help account for why the head of a relative clause cannot be questioned, clefted, etc., though other reasons involving the intimate connection between the head noun and the relativized element may ultimately be more significant.

There is a problem that forces us to revise (3.90) in light of the Pied Piping Convention. The sentence (3.211) is perfectly good, but an analysis of its islands will give the
same result we encountered when we considered the ungrammatical (3.209).

(3.211) The man whose life you saved slept.
Clearly the effect of pied piping \( NP_1 \) is to keep \( NP_1 \)-island intact, and it is this fact that makes (3.211) a good sentence as opposed to (3.209), just as it is this fact which (3.90) fails to capture due to the interference of the pied piping process.

Notice that (3.212) is allowed by (3.90) and (3.98):

(3.212) The man's life which you saved...

\[
\begin{array}{c}
\text{NP} \\
\text{NP} \quad \text{S}_1 \\
\text{the man's life} \quad \text{NP} \quad \text{VP} \\
\text{you} \quad \text{V} \quad \text{NP}_1 \\
\text{saved} \quad \text{NP} \quad \text{N} \\
\text{man} \quad \text{life}
\end{array}
\]

\( NP_1 \)-island = \{man, life\} and \( S_1 \)-island = \{you, saved, WH-\( NP_1 \), man, life\} and \( S_1 \)-island contains \( NP_1 \)-island. (3.212) and (3.211) are similar in that \( NP_1 \) is reordered intact in both cases, although only in (3.212) does (3.90) operate correctly. We must attach an ad hoc rider to (3.90) that states:

(3.213) When a WH-element is pied piped, for the purposes of determining overlapping, the WH is considered to be attached only to the ultimately preposed element.

Now an analysis of (3.211) gives \( NP_1 \)-island = \{man, life\}, but \( S_1 \)-island = \{you, saved, WH-\( NP_1 \), man, life\} and \( NP_1 \)-
island is contained in $S_1$-island.

The necessity for the extremely ad hoc (3.213) is related to the fact that a double role is assigned to the element \( \text{WH} \). On the one hand it indicates which element(s) is to be preposed; on the other hand it indicates which element is to become, morphologically, a \( \text{WH} \)-form. Of course this is precisely the kind of generalization that linguists look for and it can hardly be abandoned. On the other hand pied piping violates this generalization and it is not surprising that some further elaboration of (3.90) is necessary to take pied piping into account.

Another example of how (3.213) takes effect can be seen by examining the following two sentences.


b. The man whose brother's life you saved slept.

(3.214a)
Pied piping has taken place in (3.214a) so that $NP_2$ is ultimately preposed. According to (3.213) and (3.90) $S_1$-island = \{you, saved, $NP_1$, WH-$NP_2$, man, brother, life\}. $NP_1$-island = \{NP_2, man, brother, life\} and the islands overlap.

In (3.214b), however, where pied piping goes all the way to $NP_1$, $S_1$-island = \{you, saved, WH-$NP_1$, NP_2, man, brother, life\} and the islands do not overlap.

Here is one further example that illustrates how topicalization interacts with pied piping.


b. The man's life we finally saved.

In (3.215a), $NP_1$-island = \{the man, life\}, and $S_o$-island = \{we, finally saved, life\} and there is overlapping. In (3.215b), however, $S_o$-island = \{we, finally saved\} and the islands fail to overlap because they are disjoint, the correct result.

It appears, then, that we can dispense with the Left Branch Condition—it is really a subcase of the general
island constraint. This is important empirical evidence in favor of the island principles set forth so far.

Consider now the following string:

(3.126) Whose did you save life?

\[ S' \]
\[ NP \rightarrow whose \] \[ S_0 \rightarrow NP \rightarrow you \] \[ VP \rightarrow V \rightarrow save \] \[ NP_1 \rightarrow [NP +WH] \rightarrow life \]

\( NP_1 \)-island = \{\WH, life\}, but \( S_0 \)-island = \{you, save, life\} and there is overlapping. Note that the grammatical (3.217) is unaffected by (3.98).

(3.217) Whose life did you save?

In this case \( S_0 \)-island = \{you, save\}, life being pied piped, and the islands are disjoint.

There is a major difficulty with constituent questions, however. Recall that I insisted (in 3.2.2.4) on not relying on the syntactic analysis expressed in (3.216), but on a more semantically based analysis. Thus in (3.216), rather than \( S_0 \)-island, we should consider CQ-island: \{you, save, "someone's", life\}. This island overlaps with \( NP_1 \)-island, as we would expect. The problem arises when we consider (3.217): the islands persist in overlapping giving a false prediction of ungrammaticality.
The difficulty is that pied piping doesn't have a discernable semantic effect. Thus the CQ-island will be the same with or without pied piping. I know of no way around this problem short of insisting on the "syntactic" analysis for English Constituent Question Formation exemplified by (3.216). This analysis will not interfere with any of the conclusions we drew in the previous sections, but reliance on such a controversial and unsupported analysis weakens my hypotheses.

3.3.2 The function of an NP within a phrase marker at the time its corresponding island is defined by (3.90) apparently affects the strength of that island. Thus the following facts seem to hold: they can be viewed as island strengthening principles, though I will not state them formally.

Fact I: Subject NP's form the strongest NP-islands: 
(3.218) < (3.219).

(3.218) a. The hypotheses which support for is found in chapter three...
   b. The factory which the damage to was great...

(3.219) a. The hypotheses which I offer support for...
   b. The factory which I deplored the damage to...

In Rodman (1972b) I proposed a "Stranding Convention" which stated that prepositions couldn't be stranded in subject position. This ad hoc convention can be replaced by an (also ad hoc) principle stating that subject NP's have strengthened islands. For example, consider (3.218a):

241
NP₁-island = {the damage, to, the factory}; at the end of the S₀ cycle, S₁-island = {the damage, to, WH-factory, VP} and the islands overlap. Since NP₁ is a strong island, the grammaticality is greatly reduced.

It is important to note that if pied piping gives (3.220), there will not be overlapping islands.

(3.220) The factory the damage to which was great...
In such a case S₁-island = {WH-NP₁, the damage, to, the factory, VP} and there is containment instead of overlapping.

At the same time, however, pied piping may also give (3.221).

(3.221) The factory to which the damage was great...
In this case NP₁-island = {the damage, PP₁, to, the factory}; S₁-island = {the damage, WH-PP₁, ...} and the lower island overlaps the upper by a factor proportionate to the membership of PP₁ in NP₁-island. Now (3.221) does not seem as good as, say, (3.220); it is representative of a large number of cases--for example (3.222).
(3.222) The race of which an inquiry was held ...

The city into which a road runs south...

On the other hand, there are also a large number of cases
in which it is perfectly grammatical to pied pipe the prepo-
sition even though the NP is in subject position—contrary
to my hypotheses:

(3.223) The hypothesis for which the solution is found
in chapter three...

The door to which several keys were lost...

These examples pose a serious challenge to my claims
that NP's are islands, let alone my claim that subject NP's
are strong islands. Nonetheless I still wish to maintain my
claims, since they explain a number of phenomena quite
naturally. Hopefully some principled explanation to these
counter-examples will be uncovered by future research.

Fact II: NP objects of prepositions form stronger
islands than direct object NP's.

Thus (3.224) < (3.225).

(3.224) The city which we were aware of the
destruction of...

This is the man who we gave food to the
father of.

(3.225) The city which we observed the destruction of...

This is the man who we met the father of.

Fact III: NP objects in comparatives and equitatives
form stronger islands than other object NP's.

Thus (3.226) < (3.227).

(3.226) Where is the committee, which John is
{taller than} {as tall as } a member of, going to meet?
(3.227) This is the meeting place of the committee which John assassinated a member of.

Apart from its function in the phrase marker, it appears that semantic relations within an NP may have an effect on the strength of the corresponding island.

(i) Possessives form strong islands (3.228).\(^45\)

(ii) NP's containing agentives form strong islands (3.229).

(iii) Partitive NP's form strong islands (3.230)\(^46\)

None of (3.228)-(3.230) are as good as (3.231).

(3.228) Who did you see the father of?

John, we like the father of.

(3.229) Who did you study an offer by?

John, we studied an offer by.

(3.230) Which girls did you like each of?

These girls, we liked each of.

(3.231) What did you discover evidence for?

What did you find a key to?

These claims, we discovered evidence for.

This door, we found a key to.

Who did you see a picture of?

John, I saw a picture of.

This smattering of facts indicates the complexities and difficulties involved in a thorough description of NP-islands. What we can say is that the relative degrees of ungrammaticality, and the frequent inability to state whether a string is or is not grammatical, indicate that NP-islands are fuzzy. We are dealing with a very idiosyncratic part of
the grammar that is still poorly understood. Nonetheless (3.98) and (3.90), as modified for pied piping, are capable of handling most if not all these data once the factors that determine the strengths of NP islands are better understood.

3.3.3 In 3.1.3.1.1.2 I gave the "identifiability squish," re-exemplified here as (3.232), and noted that the lower an entry in the hierarchy the weaker its membership in its island.

(3.232) John is looking for \{some\} unicorn.
\{a\}
\{the\}
\{that\}
\{Mary's\}
\{Mary's boyfriend's\}
\{uncle's\}

This hierarchy plays a role in the determination of the strengths of NP-islands, though strangely enough it has an effect on whole islands opposite to its effect on nodes within islands. That is, an NP with a determiner chosen from the top of the hierarchy will be a stronger member within an island, but as an island itself, it will be weaker. Thus in (3.233), the strings get worse as you go from top to bottom.\(^{47}\)

(3.233) \{Who did you see\} \{a\}
\{What did you discuss\} \{some\}
\{portrait of\} \{the\}
\{proof of\} \{this\}
\{Mary's\}

Mary's uncle's.

This hierarchy is not ad hoc to the case in point. For example the rule of Genitive Head Deletion penetrates less highly identified NP's, as we would expect (cf. 3.2.2.6); so does Reflexivization: In (3.234) the strings higher up

245
are better. Similarly in (3.235).

(3.234) Here is a snapshot of Ed's lover, and over there you'll find the portrait of that Mary's Mary's uncle's the Queen's.

(3.235) I showed Mom the portrait of myself.

\{ a/some the that Mary's Mary's boyfriend's \}

We may conclude that the higher the degree of identifiability of an NP, the stronger an island that NP will form.

3.3.4 I showed in 3.3.1 how the Left Branch Condition can be replaced by a somewhat less ad hoc principle tied to the island constraints. Further support for my claim that NP's are islands, and for (3.90) and (3.98), comes from the fact that with no further statement we can delete the extremely ad hoc convention of bounding (Ross 1967:166).

(3.237) Any rule whose structural index is of the form...A Y, and whose structural changes specifies that A is to be adjoined to the right of Y, is upward bounded.

Extraposition from NP (ExNP) is "upward bounded" according to this convention: Ross (1967:Ch1) argues for the following formulation of ExNP, which we assume throughout this subsection. 48

(3.238) \textbf{Extraposition from NP:}

\[
X - \_NP[NP - S]_{NP} - Y
\]

\[
\begin{array}{ccc}
1 & 2 & 3 \\
1 & \emptyset & 3+2
\end{array}
\]

246
From underlying (3.239), then, ExNP can be applied to derive (3.240), but is blocked from deriving (3.241) by (3.237):

(3.239) That the girl that I like left bothered me.
(3.240) That the girl left that I like bothered me.
(3.241) *That the girl left bothered me that I like.

In section 2.3.2.2 we not only countered Ross's argument that ExNP is post-cyclic, but we gave three arguments in favor of its cyclicity. The third argument, taken from Chomsky (1971), is repeated here. Chomsky stated that:

(3.242) No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

Thus rules cannot "reach down" and affect solely the "contents" of some cyclic node already passed in the cycle. It is clear that this condition mitigates against the post-cyclicity of ExNP. Thus consider (3.239):

```
S_0
  \_ NP
    \_ S_1
      \_ that
          \_ NP_2
               \_ the
                   \_ N
                        \_ girl
                             \_ I like the girl
                    \_ S_2
                         \_ left

       VP
        bothered me
```

In deriving (3.240) ExNP moves S_2, as shown:
According to (3.242) ExNP can only operate on the $S_1$ cycle to derive this structure, since the effect of the transformation is confined to $S_1$.

Now to derive (3.241), ExNP would perforce operate on the $S_0$ cycle:

At the end of the $S_1$ cycle, ExNP has not taken place, so $NP_2$-island = \{the girl, $S_2$\}. At the end of the $S_0$ cycle $S_1$-island = \{the girl, left\} and these two islands overlap, predicting the ungrammaticality of (3.241).
Note that in the previous case (3.240), $\text{NP}_2$-island = \{the girl\} and $S_1$-island = \{the girl, left, that I know\} and there is containment.

The same situation obtains when the rules of Extraposition and Extraposition from PP are considered, two other rules which the ad hoc (3.237) is needed to constrain.

3.3.5 NP objects of "adverbial" prepositions such as because of, after, and in spite of form somewhat stronger islands than the NP objects of other types of prepositions, probably for the same reason that sentential objects of such prepositions form fairly strong islands. Thus (3.243)<br><br>(3.243) It was John who we left town \{because of\} a colleague of.
\{in spite of\}
\{after\}

(3.244) It was John who we \{went to town with\} \{gave a key to\} \{took ten dollars from\} a colleague of.

3.3.6 We have already discussed (in section 3.2.2.4) and illustrated how the NP dominating a series of conjoined NP's forms a strong island. Needless to say that the NP conjuncts themselves form strong islands, and that (3.90) and (3.98) account for the ungrammaticality of (3.245bc).

(3.245) a. We inspected the evidence for these claims and the solution to these problems.

b. *These are the claims that we inspected the evidence for and the solution to these problems.

c. *These are the problems that we inspected the evidence for these claims and the solution to.
Consider (3.245b):

\[ S_0 \]
\[ NP \quad VP \]
\[ these \quad V \quad NP \]
\[ are \quad the \quad N \quad S_1 \]
\[ claims \quad NP \quad VP \]
\[ we \quad V \quad NP \]
\[ inspected \quad NP_1 \quad and \quad NP_2 \]
\[ NP \quad PP \quad NP \quad PP \]
\[ evidence \quad P \quad NP \quad the \quad solution \quad P \quad NP \]
\[ for \quad the \quad claims \quad to \quad the \quad problems \]

At the end of the \( S_1 \) cycle, \( NP_1 \)-island and \( NP_2 \)-island are as shown in the tree diagram. \( S_1 \)-island, however = \{we, inspected, evidence, for, WH-claim, NP_2-island\} and overlaps with \( NP_1 \)-island. Even if pied-piping gives:

(3.246) These are the claims evidence for which we inspected and the solution to these problems.

the island constraint will still be effective, though in this case it will be \( NP_0 \)-island and \( S_1 \)-island that ultimately overlap—exactly parallel to the cases discussed in 3.2.2.4.

The explanation in 3.2.2.2.2 that we gave to account for across-the-board "violations" of the island constraint
extend to NP's. Consider (3.248):

(3.248) This is the man who the press announced
Liza's love for and Billi's devotion to.

Recall that we analyzed all "across-the-board" rule applications as involving the rule of Conjunction Reduction. Thus at the end of the S₁ cycle we get:

The islands are as shown.
At the end of the $S_0$ cycle, $S_1$-island = \{the press, announced, ... WH-man\} and NP'-island and $S_1$-island overlap. The degree of overlapping is proportionate to the degree of membership of man in NP'-island, which, because man is rightbranching in an NP object, is relatively low. (3.248) should be comparable in grammaticality to This is the man who the press announced Liza's love for. The example revolving about (3.108i) in 3.2.2.2.2 is entirely analogous to this one.

3.4 It has been my intent in this chapter to formulate a theory of islands that is at once broad enough to account for a wide range of data in several languages and simple enough to be explanatorily adequate. I have chosen to work in a framework of non-discrete or "fuzzy" grammar since it would otherwise be impossible to account for a large number of observable facts. Within the non-discrete framework it has been possible to meet the goals of generality and simplicity.

A study of Thai suggests very strongly that island constraints are not constraints on rules, but constraints on structure. We have pursued this hypothesis, and allowed ourselves to be limited by it, in formulating a theory of islands. Under this limitation there are two major goals of the theory: to define what kinds of grammatical constructs are islands, and to set forth constraints on such constructs.
In section 3.1 we hypothesized simply that all non-root S's are islands. To account for the wide ranging island strengths that are observed we further hypothesized that islands are fuzzy, and that a variety of island strengthening and island weakening factors may affect the islands in a given derivation. We noted that both syntactic and semantic phenomena could affect island strength. Resolving the interaction of the two has proven difficult, even within the "stretched" Extended Standard Theory that we have allowed ourselves, and several poorly treated problems remain. Nonetheless, the basic hypothesis, that all non-root S's are fuzzy islands, meets the criteria of generality and simplicity we aimed for.

In section 3.2 we were again able to meet the goals of generality and simplicity by positing the island constraint: islands may not overlap—which we suggest is a linguistic universal. We further undertook to determine at what point in the derivation islands are to be defined. We concluded that for English, the determination of islands is tightly linked to the notion of the transformational cycle in that island strengthening and weakening factors affect islands up until the cycle properly containing the island is passed. This means that islands are determined globally even though the island constraint itself operates on the level of surface structure.
Defined in this way, the island constraint allows us to treat the Complex NP Constraint, the Coordinate Structure Constraint, the Sentential Subject Constraint and the Left Branch Condition as manifestations of the same grammatical phenomena—overlapping islands. Moreover, with little or no additional machinery, we have been able to show that ad hoc statements regarding copying rules, across-the-board rules and bounding can be eliminated from the Grammar.

In section 3.3 we took up the hypothesis that all NP's are islands. While a number of questions remain regarding the internal relationships of non-simple NP's—in particular the roles played by prepositions and prepositional phrases—a large number of cases conformed to our hypotheses.

Since the constraints proposed in this chapter are intended to replace Ross's island constraints, I have taken care to account for most of the data he accounts for, concentrating mainly on structures that result from movement transformations since these provide the bulk of the interesting cases. Besides constraining chopping rules, Ross's island constraints acted on "feature-changing rules and unidirectional rules of deletion." We considered structures resulting from the application of deletion rules in light of our hypotheses, but failed to consider feature-changing at all. The difficulties inherent in structures affected by feature-changing rules are (1) it's unclear what features enter into—or fail to enter into—the determination of
identical elements; (2) they are among the least understood rules; (3) contrary to Ross's claims, most feature-changing rules are not constrained by his island constraints (cf. Keeld 1973). Nonetheless an extention of this study to structures affected by such feature-changing rules as Indefinite Incorporation, Reflexivization, Sequence of Tenses, etc. would certainly be a logical step.
Footnotes

Chapter Three

1 Erteschik (1972), as I understand it, is exceptional in this vein. She proposes that all S-complements are islands.

2 I will defend (3.1) most vigorously for English, though I believe it is a good candidate for a universal principle.

3 See appendix I, where the notions "position" and "type" are formally defined.

4 For now I will assume violations of island constraints as explicated by Ross (1967), and discussed in detail in chapter two, are measures of island strength valid for our discussion. Eventually I shall define, with mathematical precision, an island constraint that will incorporate these results, giving us a uniform treatment of island phenomena.

5 Again, we could be more formal (and more precise): "Given two nodes M and N belonging to the island I, where M#N, it is not necessarily the case that m_M = m_N." For the most part, such formal descriptions will be foregone when it is fairly certain that an informal statement will provide sufficient clarity.

6 Strictly speaking, we should talk about the position of a node in the phrase structure tree that corresponds to the island (since the island is an abstract, fuzzy set). A relaxation of formal terminology for the sake of readability will do no harm in this case.

7 Since my "basic" island is an S or NP, the kinds of phrase structures that are used will have some effect on my theory. There is no doubt that this hypothesis can be adjusted to any transformational grammar. In the Extended Standard Theory, for instance, I suspect the basic island would be any cyclic node. This would be extremely desirable—even elegant, for obvious reasons. The reason I continue to work in a somewhat outmoded theory is that I wish to draw on much work of my own (and others) that has been expressed in

256

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
this theory. Little solid and detailed work has been done in the Extended Standard Theory and it is not my desire to translate immense amounts of research so that it is expressible in this system, though in principle this is easy enough to do.

I have not really thought about the kinds of changes that would ensue were one to adopt a Lakovian system with its logically oriented remote structures, that is, with predicates, taking as arguments NP's and S's (prepositions). It is not implausible that islands would be defined in essentially the same way as in the present work.

8 We will discuss that-deletion and its relation to islands in section 3.1.3.2 below.

9 One version of the primacy constraint, from a lecture Ross gave in Amherst, Mass. is as follows: \((A \rightarrow B\) means "A has primacy over B")

**The Primacy Constraint**

If a rule can apply to node B, or in environment B, or from trigger B, then the rule must also be able to apply to any node A, or in any environment A, or from any trigger A, if \(A \rightarrow B\).

The problem is that subjects have primacy over objects, but chopping rules inevitably are harder to apply to subjects than objects, as we have seen.

A possible escape from this counterexample to primacy would be to adopt my hypothesis that the unmarked state of affairs is for all S's and NP's to be islands, and that when a chopping rule removes an element from an S (or an NP) there is always to some degree a violation of the island constraint. The primacy hypothesis would then state (in part) that if a fuzzy constraint applies to two nodes, it applies more heavily to the node that has primacy. Of course this statement needs empirical testing, which it will not get here. By stating, as I shall do in (3.23), that a node that has primacy has a higher degree of membership in a given island, I am achieving this very effect insofar as the island constraint is concerned.

10 I am avoiding the usage of "definite" here because there is evidence (much of it to be found in Postal (1966) that suggests that definite is a discrete notion. In the future I shall shorten the bulky phrase "identifiability of the referent of the NP" to "identifiability of the NP" where it is clear what is meant.
This is an instance of a more general situation:

\[ i \gtrapprox ii \gtrapprox iii \gtrapprox iv \gtrapprox v. \]

(i) Who did you believe a claim that Sam was suspicious of?
(ii) Who did you believe the claim that Sam was suspicious of?
(iii) Who did you believe that claim that Sam was suspicious of?
(iv) Who did you believe Bill's claim that Sam was suspicious of?
(v) Who did you believe the President's agent's claim that Sam was suspicious of?

The more highly identified the head noun in an N-S construct, the greater its degree of foregrounding, hence the greater the degree of backgrounding of the S, hence the stronger the island formed by that S, as the data suggest. We have here yet further relevance of the identifiability squish (3.24).

Though we must reject Chomsky's (1971) work because it cannot cope with non-discreteness (among other reasons), there seems to be validity to his claim that you can reorder more than one WH-constituent, which is manifested syntactically in his system by the existence of a COMP-node through which preposed WH-constituents must move (so if COMP is already occupied by some WH-constituent, another WH-constituent cannot use it). This amounts to confining the constraint (3.35) to nothing but WH-constituents, which appears to be the right sort of thing to do, since Passive and Tough-movement, two of the major counterexamples to (3.35) (cf. 2.3.1), do not involve WH-constituents exclusively.

See for example Greenbaum & Quirk (1970) and Quirk & Svartvik (1966).

The irony is that it is just this inability to get "reliable" data in a discrete grammar that suggests so strongly that grammars are fuzzy. An interesting exercise is to look at a copy of "Some Classes of English Verbs" by D. Alexander and W.J. Kunz, or "Adjectives Before That-Clauses in English" by D. Alexander and P.H. Matthews, both 1964 publications of the Linguistics Research Project of Indiana University, available from the Indiana University Linguistics Circle. There are plenty of verbs and adjectives in these many pages of data to try out "islandness" on, and they are subcategorized in numerous ways so that classes of similar predicates are available. Choose at random six or eight predicates which take that-S complements and determine (if you can) the relative island strengths of their complements. Now let a colleague do the same thing, or you yourself try it after a week has elapsed. The results are frequently inconsistent.

The significance of these remarks in this work on fuzzy grammar should be evident. There are myriad data which lead to hopeless and genuine confusion when we try to
determine relative grammaticality. I claim this is a result of our attempt to discretize non-discrete data with intuitions that are incapable of such fine measurements. As noted in chapter one, intuitions are subject to a kind of uncertainty principle in that a penetrating consideration of some datum is apt to affect the intuition. Thus our (mostly subconscious) grammar makes finer distinctions than our (mostly conscious) intuitions are capable of fathoming. See Labov "Negative Attraction and Negative Concord" in Language 48.4, especially p. 777 for further remarks in this vein.

15 But Lakoff (1973) has observed hierarchies similar to (3.53) which usually interact with syntactic processes in the expected linear way, but in certain special cases interact in such a way as to "map the...hierarchy into a sort of parabolic curve." (p. 12).

16 This string seems a little worse because it is also presupposed. There is, actually, a small conspiracy to make complements of factive predicates somewhat stronger islands. See footnote 17 below for comment on (3.57c), which is slightly more grammatical than we would predict.

17 I'm not sure where for-to complements fit in the clausiness squish. Most likely between tenseless that-clauses and subjected non-finite clauses. For-to clauses are acted upon by opposing forces as far as their island strengths are concerned. On the one hand, they are nounier, increasing their island strength, but on the other hand, they are less clausy, reducing their island strength. This is somewhat in evidence in (3.57), where the e-sentence is slightly better than expected. For further discussion see 3.2.2.8.

18 To say this matter is "simple" is misleading. In fact I can easily envision a model of Transformational Grammar that incorporates into semantic representations whatever information is reflected syntactically by such phenomena as Topicalization, Passive, Dative Movement, Extrapolation, and so on. It would then be feasible to define everything on a/the level of semantic representation, although Sluicing (discussed in 3.2.2.6) would have to be accounted for. We shall nonetheless push our syntactic, surfacly approach as being most interesting at this stage of our knowledge.
I am being imprecise. The islands discussed here and below really consist of nodes, not the lexical items dominated by the nodes. Nothing relevant to the theory is lost by this convenient oversimplification, and much is gained in the way of clarity. Also, we avoid delicate matters such as the exact form of derived constituent structures.

The S cycle ends when the last cyclic or last "last cyclic" rule is applied (if there are "last cyclic" rules). Postcyclic rules may still apply, however. The post-cyclic rules are those rules that can "reach down" indefinitely far into the tree. They include the WH-Preposing rules (cf. Postal (1971) for some discussion), Topicalization, Dislocations (cf. Emonds (1970) for discussion), and Pronominalization.

In dialects like standard English, where sentences such as (3.89) are of low grammaticality, other factors than principle (3.80) are responsible.

Any principle of pruning will be important to these considerations. However, in light of the evidence in favor of fuzzy category nodes (cf. Ross (1972)), the whole meaning of pruning is altered since pruning is a relic of discrete grammars. The entire subject needs to be looked into again from the point of view of fuzzy grammar. In 3.3 we touch upon a single aspect of this problem.

One way to do this is to take the ratio of the number of elements involved in overlapping to the number of elements not involved. The closer to unity, the higher the degree of overlapping.

I am oversimplifying slightly. We would also have to take the degree of intersection into account for islands that are nearly, but not quite disjoint. Nothing is gained, however, by spelling out the details of how to precisely measure degree of intersection and relate it to degree of overlapping, which could be done easily.

This intuition comes across better if one tries to question the position occupied by the guard:
(i) Who do you deplore that John believed that Mark murdered?
(ii) Who do you believe that John deplores that Mark murdered?
The second of these seems to be slightly worse, just as its counterpart (3.97). This intuition probably occurs because the preposed NP in (ii), being embedded directly under the factive predicate, is felt to be more "frozen" in place than when its embedded position is separated from the factive predicate by an intervening verb. Principle (3.95) captures this fact.

26 A slight complication in collapsing the two rules occurs because certain prepositional phrases can be topicalized but not left-dislocated
(i) With an axe, she gave her mother forty wacks.
(ii) With an axe, she gave her mother forty wacks#???
It seems likely that a general principle of grammar will prohibit pronoun-leaving rules from operating when the grammar fails to provide a pronominal form.

27 This does not seem to involve a relative construction. I find (i) bad, but (ii) fine:
(i) Maxwell isn't half the doctor which his father is.
(ii) Maxwell isn't half the doctor that his father is.
This fact is reflected in the structure I give, which otherwise need not be taken too seriously.

28 (From Ross (1967): ex 6.58) Conjunction Reduction

a. \[
\begin{array}{c}
\text{and} - \{X - A\}^n_B \\
1 & 2 & 3 \\
(1 & 2 & 0)_B
\end{array}
\]
\[
\text{OPT}
\]

b. \[
\begin{array}{c}
\text{and} - \{A - X\}^n_B \\
1 & 2 & 3 \\
(1 & 0 & 3)_B
\end{array}
\]
\[
\text{OPT}
\]
Condition: all occurrences of A are identical.

29 Ross calls these "niches," and has discussed the problem of determining where in a sentence a niche is allowed in his unpublished paper "niching."

30 It seems to me, however, that if the that is deleted, pauses judiciously inserted, and lowered intonation placed on we believe, a parenthetical reading is still possible:

(i) The lock of hair - PAUSE - which - PAUSE - we believe - PAUSE - Ronald sent Spiro, came from the private collection of the Lieutenant Governor.
I cannot presently explain this oddity.
I regret that John has to quit school, since he has no money violates the root-S constraint. As, since, although, for unambiguously introduce parenthetical clauses.

This argument is slightly weakened by the fact that there are non-restrictives that do not have a fully grammatical parenthetical paraphrase:

?We rarely go to massage parlors, and they are very stimulating.

We rarely go to massage parlors, which are very stimulating.

The starred strings are possible if the non-restrictive reading is on the entire string, as we would, of course, expect.

The fact that non-restrictive clauses ultimately have a reordered element also contributes to their strength as islands. This was discussed in 2.3.1 and 3.1.3.

More exactly, (or less exactly as it were) the island constraint operates on whatever formal entities represent the steps of a derivation, taking into account both syntactic and semantic information. It is questionable that trees will be the sole representative device for all this information.

I have no basis for deciding whether "someone" and "she" are to be considered different elements for the purposes of determining overlapping— I'm ignoring this minor problem.

Very heuristically, I believe that pregenerated structures in general, and echo-questions in particular, are generated with a reduced amount of semantic structure and that fact, at least in part, accounts for the possibility of violating constraints like (3.98) that are based partly on semantic structure. Van Lancker (1972) has claimed that propositional modes of language use are lateralized to the left cerebral hemisphere, while automatic modes (such as swearing, emotional expletives, etc.) are bilaterally represented in the brain. Though she doesn't mention echo-questions per se, there is no doubt that there is some
automaticity in producing echo-questions. This suggests that the left cerebral hemisphere, possibly the source of constraints like (3.98), is not as involved in the production of echo questions, or other kinds of pregenerated structure, as it is in "normal" speech, hence providing an explanation of why such constraints are relaxed in certain circumstances.

38 The rule involved here is worthy of study in its own right. There are a number of restrictions. For example I find (i) > (ii) > (iii) > (iv): 

(i) Fluffy is sick, which \{few people\} realize(s).
\{no one\}

(ii) Fluffy is sick, which \{everyone\} realizes.
\{he\}

(iii) Fluffy is sick, which \{few people\} believe(s).
\{no one\}

(iv) Fluffy is sick, which \{everyone\} believes.
\{he\}

The clause containing the relativized sentential is best as a negated factive, okay as a factive, and not so great as a non-factive.

The interaction with factivity is not surprising since the relativized sentence is "definite," that is, its reference is known or "presupposed" so it is most compatible with a factive predicate. I don't understand why negatives are better, if indeed they are: Is (v) better than (vi)?

(v) Fluffy is sick, which doesn't surprise me.

(vi) Fluffy is sick, which surprises me.

The actual analysis of relativized sentential will not influence the conclusions drawn with respect to island constraints.

39 This is one of those interesting facts that lends credence to Lakoff's (1970) claims that there are no VP's in DS, only sentences, predicates and nominals. VP's he claims, are the surface structure remnants of sentences.

40 The pseudo-cleft for some reason is better than the cleft. It was to write a book that I believed that John decided to try. Emonds (1970) accounts for this fact as part of the structure preserving hypothesis. His theories disallow base structures such as V-VP-S, and hence forbid dcs's of chat form. But V-VP-S is exactly the structure required for a clefted VP. See tree (3.161c).

263
The parenthetical remark is because Postal considers "PP's" to be NP's.

This sentence has a reading relative to which it is grammatical, and which corresponds to the following derivation: The claim that W.C. Fields was an oblate spheroid, which we debated, can never be proven for sure—(by Extraposition from NP) The claim which we debated that W.C.Fields was an oblate spheroid can never be proven for sure = (3.210).

This "ad hoc" principle has a chance to become well motivated—the Stranding Convention does not. There is certainly something about being in subject position that affects islands and the elements of islands. I have attempted to explain what that "something" might be in 3.1.3.1.1 and 3.2.2.7. In the case at hand, however, I am at a loss for a plausible explanation.

Strings such as The dog the man with which was wearing a hat bit me are decidedly weird—but for reasons having to do with constraints on pied piping.

But This is the boy that he is the father of seems mysteriously, to be of high grammaticality.

For an alternate explanation to this fact, see Rodman (1972b:99f).

Some seems to follow a in this case for reasons not clear to me.

We also assume Extraposition and Extraposition from PP are similarly formulated.

The SD's of all these rules have a variable, as (3.238) shows. Postal (1971) distinguishes between "essential" and "non-essential" variables. Essentially, a non-essential variable is an "abbreviating variable," without which the rule could be formulated by means of a disjunction in the SD. An essential variable is one without which the rule could not possible be formulated (e.g. WH-Q). The variables in the "extraposition" rules are non-essential and the addition of this distinction to the grammar of English may also allow bounding to be dispensed with, and account for why elements cannot be "beyond the first sentence up."
Chapter Four

No thesis is complete without at least a brief discussion of the implications that the work has for the theory of grammar. In our case the most obvious and significant implication is that grammars are fuzzy. In particular, islands are fuzzy, and the island constraint, as a result, is also fuzzy.

We have also suggested that the island constraint, once thought to operate on rules, actually operates on linguistic structures. The determination of islands must be based on the linguistic cycle when such an approach is taken, so our work supports the existence of that notion. Islands also need to be defined transderivationally under this basic assumption. Even when island constraints were confined to constraints on rules, the fact that clauses which were extraposed out of complex NP's retained their status as islands, and the facts revolving around the rule of Sluicing, suggested that the Complex NP Constraint was a global constraint. By conceding the globality of the island constraint and making it in effect a well-formedness constraint on grammatical structures, we have been able to augment the observational adequacy of the constraint both in its universal applications and its language specific applications.
We also noted that both syntactic and semantic factors affect the strength of islands, and that if islands were defined from strictly syntactic information it would not be possible to treat certain aspects of islands adequately. We must either "pack" more semantic information into the syntactic derivation, or concede that there are other components in the grammar which interact with the component that produces surface structures.

We suggested on several occasions that fuzzy category nodes are not only consistent with our theory of fuzzy islands, but are necessary. One particular instance—the node dominating complements—is discussed in the Appendix where it is argued that this node is partly noun phrase, partly sentence. The formalism for making the notion of fuzzy category node explicit is also given there. Ross (1972) has other evidence that suggests the need for fuzzy category nodes in the grammar.

Fuzzy rules are undoubtedly needed in any adequate fuzzy grammar. There are two types of fuzzy rules: fuzzy obligatory rules and fuzzy optional rules. A fuzzy optional rule is a rule whose application, under certain circumstances, results in decreased grammaticalness; a fuzzy obligatory rule is a rule whose failure to apply under certain circumstances, even though its structural description is met, results in decreased grammaticalness. The reason two types of fuzzy rules are needed is that in some cases it is the
actual application of a rule that lowers grammaticality, and in other cases it is the non-application.

Particle Movement is an example of a rule that is simultaneously fuzzy optional and fuzzy obligatory. It is a fuzzy optional rule relative to the nouniness squish (3.53)---as you attempt to move particles over structures that are higher in the squish, the results get successively worse and worse. That is, \(a \leq b \leq c \leq d \leq e \leq f \leq g\) (cf. (3.55), reproduced here as (4.1)).

(4.1) a. I want you to write that linguists are having trouble finding jobs down.

b. Would you please call for them to dispatch a cab up.

c. Let's look what they bought at the candy store over.

d. The group wanted to talk his explaining the problem over.

e. The coed brought her teacher's explaining of the problem up in class.

f. The clerk wrote their explanation of the problem down.

g. A teacher shouldn't put her students down.

On the other hand, Particle Movement is a fuzzy obligatory rule in that its failure to operate under certain circumstances produces sentences of lowered, but not necessarily zero, grammaticality. The sentences (4.2), in which Particle Movement has not been applied, never decrease in grammaticality as you go from top to bottom.\(^1\)

267

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(4.2) a. Mary called up him.
   b. Mary looked up herself in the directory.
   c. You students ought to write down this.
   d. Did you call up \{someone.\}
   \{anyone.\}
   e. David called up a hooker and Ricky called up one too.
   f. The secretary calls up the bookstore on Mondays.

The distinction between fuzzy obligatory and fuzzy optional is necessary in a fuzzy transformational grammar, but what about a fuzzy well-formedness grammar such as the kind advocated by Lakoff? In a non-fuzzy well-formedness grammar the distinction between obligatory and optional is reflected in the fact that obligatory rules have a blocking condition: if the "input" structure meets Structural Description I and the "output" structure fails to meet Structural Description II, then the derivation blocks. Optional rules lack such a condition.

If, in a fuzzy well-formedness grammar, the distinction between fuzzy optional rules and fuzzy obligatory rules were not necessary, then this fact would favor the well-formedness grammar over the transformational grammar. The distinction, however, must be maintained in spite of the fact that in a fuzzy well-formedness grammar all rules have blocking conditions that indicate ill-formedness to a degree. With fuzzy obligatory rules, if the degree of blocking is 1.0, the rule acts like an ordinary obligatory rule in that the failure of the "output" structure to meet Structural Description II

268
results in total ungrammaticality. As the degree of blocking becomes smaller and smaller, the failure of the output structure to meet Structural Description II results in less and less ungrammaticality. At a value of 0.0, the rule would behave like an ordinary optional rule.

Fuzzy optional rules work the same way, but with an important difference—the difference that maintains the distinction between the two kinds of rules: if the degree of blocking is 1.0, then the success of the output structure in meeting Structural Description II results in total ungrammaticality (the "blocks" in discrete transformational grammar). As the degree of blocking decreases, the success of the output structure in meeting Structural Description II results in less and less ungrammaticality. At a value of 0.0, of course, the rule would behave like an ordinary optional rule.

We conclude, then, that the distinction between fuzzy optional and fuzzy obligatory rules is needed in both fuzzy transformational grammar and fuzzy well-formedness grammar, and fuzzy rules cannot help us choose between the two opposing theories.

Returning to transformational grammar, it might be suggested that rules are not classified as obligatory or optional, but rather are assigned a "coefficient of applicability" that ranges as follows:
1.0 Obligatory
   . Failure to apply
   . results in less
   . and less loss of
   . grammaticality
0.0 Optional
   . Application results
   . in more and more
   . loss of grammaticality
-1.0 Blocks

The coefficient of applicability would, of course, be determined by, and interact with, other grammatical entities (e.g., the nouniness squish in the case of Particle Movement).

While this appears to be a generalization not capturable in a well-formedness grammar, it is merely a clever notational variant of the fuzzy optional/fuzzy obligatory distinction and does not have any deeper significance.

The notion of "preferred" reading can be captured in a satisfying way in a fuzzy grammar. One posits that a sentence has a degree of grammaticality relative to each of its semantic readings. Reading A is preferred over Reading B if the coefficient of grammaticalness of the sentence relative to Reading A exceeds the coefficient relative to Reading B. For example, consider the following sentence: (From Jackendoff (1971:495)).

(4.3) It is possible that John caught a fish.

270
This sentence, for many persons, can be construed ambiguously. In one reading "a fish" is non-specific; in the other reading "a fish" is specific. The preferred reading, however, is the non-specific one. The specific reading is less preferred because relative to this reading, the sentence is not fully grammatical, as (4.4) suggests.

(4.4) ??It is possible that John caught a fish, because he ate it.

Suppose we are working in a theory of grammar in which the semantic representation of a sentence and its deep structure are the same abstract entity. Then it follows from the assertion that the coefficient of grammaticalness depends on the whole derivation that each reading of an ambiguous sentence will have its own coefficient of grammaticalness. In a theory that relies on rules of semantic interpretation, if we are to maintain our hypothesis about preferred readings, we will have to allow for the possible existence of fuzzy rules of semantic interpretation that have the potential of lowering grammaticalness relative to a particular reading. Although the former theory may seem less messy in this one regard, it is not clear whether either theory gains an edge from this one consideration because the theory of preferred readings is untested.

An interesting and related consideration leads to what I call the "Maximal Coefficient Principle." This principle states that if a surface structure relative to a particular reading has two or more derivations, the coefficient of
grammaticalness will be the maximum from among the possibilities.

An application of the Maximal Coefficient Principle is needed to account for the high grammaticality of (4.5).

(4.5) The man who I read a statement about has resigned.

In Rodman (1971) I argued that (4.5) has two possible underlying sources corresponding to these two underlying relative clauses.

(4.6) a. 

\[
\begin{array}{c}
S \\
\downarrow \\
NP \\
I \\
V \\
read \\
NP \\
a \text{statement} \\
NP \\
S \\
VP \\
\end{array}
\]

b. 

\[
\begin{array}{c}
S \\
\downarrow \\
NP \\
I \\
V \\
read \\
NP \\
a \text{statement} \\
NP \\
PP \\
\end{array}
\]

Sentence (4.5) is unambiguous, just as (4.7) is unambiguous.

(4.7) I read a statement about that man.
Yet in (4.5) and (4.7) there are two distinct underlying structures, and two distinct derivations. In the case of (4.5) there are, in fact, two different coefficients of grammaticalness since there is a violation of the island constraint if (4.6a) is in the underlying structure of (4.5). The Maximal Coefficient Principle, however, tells us that speakers will construe (4.5) as being derived with (4.6b) in the underlying structure.

Note that since (4.8) is a normal passive sentence, structures such as (4.6a) must be allowed.

(4.8) A statement about that man was read by me.

Of course the grammaticality of (4.5) speaks for the need for (4.6b). Further evidence that both (4.6a) and (4.6b) are needed in the underlying structures of (4.7) is given in Rodman (1971).

Further evidence in favor of the Maximal Coefficient Principle is given by Postal (1971:158f). He notes that (4.9) is grammatical and may be derived either by a direct application of Y-movement to the NP to him₁, or by an application of Y-movement to him₁ in conjunction with pied piping.

(4.9) To him₁, the man₁ claimed you were engaged.

Significantly, the Crossover principle would block (4.9) if it came from an application of Y-movement to him₁ with pied piping. Postal states however:

But is is a general principle that if a structure has a valid derivation, this is what we perceive, regardless of the number of ill-formed derivations which it has. That is, something is a sentence

273
with the structure K if it has a valid derivation in the grammar which assigns it K. And under these conditions, we pay no attention to the fact that it failed to be derived in n other ways. This point can be seen in many of our other examples.

In terms of fuzzy grammar, this "general principle" is, of course, our Maximal Coefficient Principle.

A theory of fuzzy grammar will have far-reaching effects on our understanding of the nature of human language. I hope that my study of fuzzy islands will ultimately contribute to the development of such a theory.
Footnotes

Chapter Four

These data imply the existence of a pronoun hierarchy.

Definite Pronoun
Reflexive Pronoun
Demonstrative Pronoun
Indefinite Pronoun
Noun

This hierarchy interacts with at least one other rule, namely Dative Movement, which is a fuzzy optional rule relative to it.

Additional details were given in Rodman (1972).

The blocking conditions are determined by other factors in the grammar.
BIBLIOGRAPHY


____. 1968. MS. Remarks on Nominalization.

____. 1971-72. MS. Conditions on Transformations.

Culicover, P. 1967. MS. The treatment of idioms within a transformational framework.


Keenan, E. 1973. Logical expressive power and syntactic variation in natural languages. Paper delivered to Department of Linguistics, UCLA.


Labov, W. Negative attraction and negative concord. Language 48,4.


Ross. 1967. MS. Constraints on variables in syntax.


Appendix

In the course of this work we have hypothesized that the position of a node in a tree may affect the strength of the node in its island. We also have talked about fuzzy nodes, that is, nodes which do not belong exclusively to one grammatical category or another, but which behave as if they belonged simultaneously to two or more categories, exhibiting different degrees of "strength" relative to the respective categories. These notions can and should be made formally precise, and the intention of this Appendix is to do so. We shall draw on section II of McCawley's 1968 paper "Concerning the Base Component of a Transformational Grammar" to achieve these ends.

McCawley gave an axiomatic definition of "the mathematical object known as a tree (more precisely, a rooted, labeled, oriented tree; ...)." His definition is an excellent jumping off point for our own work, and I reproduce it here, with a couple of irrelevant notational changes:

A tree is a finite set of objects (called 'nodes') with three relationships D 'directly dominates', L 'is to the left of', and B 'bears the label', satisfying the following axioms:

1) there is a node $x_0$ such that for no node $x$

$$x D x_0$$

($x_0$ is called the 'root' of the tree);
(2) if \( x \) is a node distinct from \( x_0 \), then \( x_0D^x_1 \), where \( D^x_1 \) is the relationship that holds between two nodes \( a \) and \( b \) if there is a chain of nodes \( a_1, \ldots, a_n \) such that \( aDa_1, a_1Da_2, \ldots, a_nDb \) (\( D^x_1 \) can be read 'dominates'; this axiom asserts that a tree is 'connected');

(3) if \( xDy \) and \( x'Dy \), then \( x=x' \) (i.e. a tree contains no 'loops');

(4) \( L \) is a partial ordering on the nodes (i.e. if \( xLy \) and \( yLz \), then \( xLz \); if \( xLy \), then it is false that \( yLx \);

(5) for any two nodes \( x \) and \( y \), if \( x\neq y \), then either \( xD^x_1y \) or \( yD^x_1x \) or \( xLy \) or \( yLx \);

(6) if \( x \) is non-terminal (i.e. if there is a \( z \) such that \( xDz \) and \( xLy \), then there is an \( x' \) with \( xDx' \) and \( x'Lx' \); if \( y \) is non-terminal and \( xLy \), then there is a \( y' \) with \( yDy' \) and \( xLy' \);

(7) every node bears the relation \( B \) to exactly one element (its 'label'), the possible labels being a set of objects distinct from the nodes;

(8) if \( wD^x_1x \), \( wD^x_1z \), \( xLy \), and \( yLz \) then \( wD^x_1y \); a tree not possessing this property would be 'discontinuous': there would be a node \( w \) which dominated two nodes \( x \) and \( z \) without dominating all the nodes that were between them.

The eight axioms are not independent. Axioms (6) and (8) can be derived from axioms (1)-(5), as we now prove:
DEPENDENCE OF AXIOM 6:

Part I

Given:  \( x \) is non-terminal; \( xLy \)

To Prove: There is an \( x' \) with \( xDx' \) and \( x'Ly \).

Proof: Since \( x \) is non-terminal there exists a node \( x' \) such that \( xDx' \), by definition of "terminal."

By axiom (5), one of the following must hold:

(i) \( x'D^*y \)
(ii) \( yD^*x' \)
(iii) \( x'Ly \)
(iv) \( yLx' \)

We shall proceed to show that the assumption of (i), (ii), or (iv) leads to a contradiction, and that therefore (iii) must hold, completing the proof.

Contradiction from (i): \( x'D^*y \)

To derive the required contradiction, it is necessary to prove the following lemma:

Lemma: \( D^* \) is a transitive relation.

Proof: Let \( a, b \) and \( c \) be distinct nodes such that \( aD^*b \) and \( bD^*c \). We must prove \( aD^*c \). By axiom (2) there is a chain of nodes \( x_1, x_2, \ldots x_n \) such that \( aDx_1, x_1Dx_2, \ldots x_nDb \), and a chain of nodes \( y_1, \ldots y_m \) such that \( bDy_1, y_1Dy_2, \ldots y_mDc \).

From this it follows that there is a chain of nodes \( x_1, \ldots x_n, b, y_1, \ldots y_m \) such that \( aDx_1, \ldots x_nDb, bDy_1, \ldots y_mDc \), which, by axiom (2), means that \( aD^*c \). QED (Lemma)
Now we know that xDx', so xD*x' too. (We allow the "empty" chain in axiom (2), in other words.) If x'D*y were to hold, then by the Lemma, xD*y would hold. But we are given xLy and so xD*y is impossible by axiom (5). QED (i)

Contradiction from (ii): yD*x'

Suppose that yD*x'. Then, by axiom (2) there is a chain of nodes a₁, ... aₙ such that yDa₁, a₁Da₂, ... aₙDx'. We saw that xDx', so by axiom (3) x=aₙ. But now there is a chain a₁, ... aₙ₋₁ such that yDa₁, a₁Da₂, ... aₙ₋₁Dx, and by axiom (2) we have yD*x. But this is a contradiction to the given fact that xLy, by axiom (5). QED (ii).

Contradiction from (iv): yLx'

Since we are given xLy, and axiom (4) states that L is a partial ordering (which by definition means that L is transitive), by assuming that yLx' we may conclude xLx'. However xD*x' is known to hold, so by axiom (5) xLx' is impossible, and a contradiction is reached. QED (iv).

Only (iii) can hold, then, and the proof of Part (I) of axiom (6) is completed.

Part II

Given: y is non-terminal; xLy

To Prove: there is a y' with yDy' and xLy'.

Proof: Since y is non-terminal there exists a node y' such that yDy'.
By axiom (5), one of the following must hold:

(i) $xD^*y'$
(ii) $y'D^*x$
(iii) $xLy'$
(iv) $y'Lx$

We proceed as in Part (I).

**Contradiction from (i): $xD^*y'$**

The proof is analogous to case (ii), Part (I). If $xD^*y'$ then there exists $a_1, \ldots, a_n$ such that $xDA_1, \ldots, a_nDy'$. Since $yDy'$ implies, by axiom (3), that $y=a_n$, we have a chain $a_1, \ldots, a_{n-1}$ such that $xDA_1, a_1Da_2, a_{n-1}Dy$. This means that $xD^*y$, a contradiction to $xLy$. QED (i).

**Contradiction from (ii): $y'D^*x$**

We have seen that $yD^*y'$. Now $yD^*y'$ and $y'D^*x$ would imply $yD^*x$, if we assumed $y'D^*x$. This contradicts $xLy$. QED (ii).

**Contradiction from (iv): $y'Lx$**

Assuming $y'Lx$ and noting the given fact $xLy$, we get $y'Ly$ by the transitivity of $L$. But $yDy'$, so $y'Ly$ is impossible by axiom (5). QED (iv).

This completes the proof of Part (II), and hence the proof of the entire axiom.

**DEPENDENCE OF AXIOM 8:**

**Given:** $wD^*x, wD^*z, xLy, yLz$

**To Prove:** $wD^*y$

**Proof:** By axiom (5), one of the following four situations must obtain:
(i) \( yD^*w \)
(ii) \( xL^*y \)
(iii) \( wL^*y \)
(iv) \( yLw \)

Strategy: eliminate (ii), (iii) and (iv):

**Elimination of (ii)**
Assume \( yD^*w \). Then by the transitivity of \( D^* \) and the given fact \( wD^*x \) we may conclude \( yD^*x \), a contradiction to axiom (5) since we are given \( xLy \). QED (ii).

**Elimination of (iii)**
Assume \( wL^*y \). Then by the transitivity of \( L \) and the given fact \( yLz \), we may conclude \( wLz \), a contradiction to axiom (5) since we are given \( wD^*z \). QED (iii).

**Elimination of (iv)**
Assume \( yLw \). Then from \( xLy \) we may conclude \( xLw \), a contradiction to the given \( wD^*x \). QED (iv).

The elimination of (ii), (iii), and (iv) allows us to conclude (i) by axiom (5). QED.

We may now define in precise terms the following four relationships among the nodes of a tree, thus justifying our use of them in the body of the text: (i) leftmost
(ii) left of (iii) rightmost (iv) higher.

(i): Here the notion is actually "leftmost under a given node." When we say "leftmost in a tree" we mean "leftmost under the root." To define this notion precisely we draw on the notion "terminal." Only a terminal element
can qualify as a candidate for leftmost element under a node. An element $a$ is leftmost under $N$ just in case (1) $a$ is terminal and (2) for every $x$ dominated by $N$ (ND$x$x) such that $x$ is a terminal element (Tx) not identical to $a$, aLx holds. I.e., $a$ is leftmost under $N$ whenever (ND$x$x&Tx&a#x) is a sufficient condition for aLx for all $x$.

(ii): $a$ is "lefter" than $b$ under a node $N$ if ND$a$ and ND$b$ and aLb.

(iii): We mean "rightmost under a given node." An element $a$ is rightmost under $N$ just in case (1) $a$ is terminal and (2) for every $x$ dominated by $N$ such that $x$ is a terminal element not identical to $a$, xLa holds. I.e., $a$ is rightmost under $N$ whenever (ND$x$x&Tx&a#x) is a sufficient condition for xLa for all $x$.

(iv): We mean "higher within a tree." Given two nodes $a$ and $b$, there exist chains $a_1$, ..., $a_n$ and $b_1$, ..., $b_m$ such that $x_0Da_1$, $a_1Da_2$, ..., $a_nDa$ and $x_0Db_1$, $b_1Db_2$, ..., $b_mD_b$ (recall $x_0$ is the root). If $n < m$, $a$ is higher than $b$. If $n > m$, $b$ is higher than $a$. If $n=m$ the nodes are of equal height. The indices $m$ and $n$ allow us to compare heights of elements in different trees, as well as in the same tree.

To define the notion "correspond in position and type" (cf. Ex. (3.4)) we need the auxiliary notions "immediately to the left of" and "immediately to the right of": We say that $a$ is immediately to the left of $b$ iff:
(i) $a$ and $b$ are immediately dominated by the same node $N$
(ii) $b$ is not leftmost under $N$
(iii) $aLb$
(iv) for all $x$, $x\neq a$, immediately dominated by $N$ such that $xLb$, $xLa$.

NB: If a node is leftmost, no node is immediately to its left.

We say that $a$ is immediately to the right of $b$ iff:

(i) $a$ and $b$ are immediately dominated by the same node $N$
(ii) $b$ is not rightmost under $N$
(iii) $bLa$
(iv) for all $x$, $x\neq a$, immediately dominated by $N$ such that $bLx$, $aLx$.

NB: If a node is rightmost, no element is immediately to its right.

We say that two nodes $a$ and $b$ correspond in position and type iff (1) $a$ and $b$ bear the same label and (2):

a. The node that immediately dominates $a$ bears the same label as the node that immediately dominates $b$ (or no node dominates either, i.e., they are both roots).

b. If there is a node immediately to the left of $a$, then there is a node immediately to the left of $b$ and both such nodes bear the same label. However if $a$ is leftmost, $b$ is leftmost.

c. If there is a node immediately to the right of $a$, then there is a node immediately to the right of $b$.
and both such nodes bear the same label. However if a is rightmost, b is rightmost.

d. (i) The node that immediately dominates a and the node that immediately dominates b, if any, correspond in position, as defined in (a)-(d).

(ii) The node immediately to the left of a and the node immediately to the left of b, if any, correspond in position, as defined in (a)-(d).

(iii) The node immediately to the right of a and the node immediately to the right of b, if any, correspond in position, as defined in (a)-(d).

Intuitively, this complicated definition of "correspond in position and type" means that the two nodes occupy identical positions in the tree with respect to dominating nodes, sister nodes, and nodes dominating sister nodes; nodes dominated by the nodes in question, or nodes dominated by their sister nodes, are discounted.

In the two trees given below, the subject NP's correspond in position and (of course) type; the NP's in the verb phrases do not correspond in position.

```
S
   NP
     S
       That Nixon lied and lied
   VP
     V
       proved
     PP
       to the world
     NP
       his total baseness
```
The definition of fuzzy node revolved about axiom (7), which we reproduce:

**Axiom (7):** Every node bears the relation $B$ to exactly one element (its "label"), the possible labels being a set of objects distinct from the nodes.

We also need to define the notion "fuzzy relation." This in turn requires us to define the notion "relation" in set theoretic terms. A relation $R$ is a subset of the Cartesian product of two (not necessarily different) sets. For example, given the set $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$ the Cartesian product $A \times B$ is the set of ordered pairs $\{(a_1, b_1), (a_1, b_2), \ldots, (a_n, b_1), (a_n, b_2), \ldots, (a_n, b_m)\}$. The relation $R \subseteq A \times B$ and contains exactly those ordered pairs $(a_i, b_j)$ such that $a_i R b_j$ holds.

Even if neither $A$ nor $B$ is a fuzzy set (or more precisely all membership coefficients of elements in $A$ and $B$ are 1.0 or 0.0), we can nonetheless define $R$ to be a fuzzy set by assigning membership coefficients to elements of $A \times B$. Thus $(a_i, b_j, n)$ belongs to $R$ to degree $n$ ($0 \leq n \leq 1$). Alternatively, $a_i R b_j$ to degree $n$. That is, the relationship $R$ holds between $a_i$ and $b_j$ to degree $n$. If $n=0$, $R$ fails to hold;
if n=1, the relationship holds in the "usual" sense; for
0 < n < 1 the relationship holds to a certain degree measured
by the magnitude of n.

Consider axiom (7). It tells us that from a set of
nodes \{x_1, \ldots, x_n\} and a set of labels \{a_1, \ldots, a_m\} there
is a relation B ("bears the label") such that to each \(x_i\),
\(x_i B a_j\) holds for exactly one \(j\).

In fuzzy terms we might redefine axiom (7):

**Axiom (7\*)**: Every node bears the fuzzy relation B to
one or more labels, the possible labels being a set of
objects distinct from the nodes. The sum of the coeffi-
cients of relationship between a given node and its
labels must equal 1.0.

Thus the node \(x_i\) may bear the relationship B to \(a_j\) to
degree 0.5, to \(a_k\) to degree 0.3, and to \(a_r\) to degree 0.2.
Or, of course, it may bear B to \(a_j\) to degree 1.0.

Since the relation B is a categorizing relationship,
we can cut a few corners and say that \(x_i\) belongs to category
\(a_j\) to degree \(n\), and to category \(a_k\) to degree \(m\), and so on.

In particular, we will allow ourselves to say, for
example, that a node is an NP to such and such a degree and
an S to such and such a degree.

Whether or not fuzzy nodes are truly needed in the
theory of grammar is a vital open question to which little
serious attention has been paid. Ross (1972) and ("nouniness")
has discussed the subject with only a modicum of thoroughness.
A good candidate for a fuzzy node is the "sentential" complement. It is often depicted as the construct $NP[S]$, a kind of schizophrenic representation that comes from the fact that complements display the behavioral patterns of both noun phrases and sentences. An alternative solution is to posit a single node in such cases that has, to a certain degree the properties of a noun phrase, and to a certain degree the properties of a sentence. We have seen that such an analysis is easily formalized. The remaining question is whether the additional theoretical machinery is justified. The answer, I believe, is affirmative on the grounds of observational adequacy. Certain phenomena cannot be adequately described by the mere presence of $NP[S]$. We shall discuss three of these, noting that a complete investigation is far beyond the scope of this appendix.

Consider first the IISNP, discussed at some length in section 3.2.2.7. One of the clearest grammaticality squishes is that given as (3.175), where $a \leq b \leq c \leq d \leq e \leq f \leq g$, corresponding to the nouniness hierarchy:

\begin{align*}
(A1) \text{ that-S (non-factive)} \\
& \text{ that-S (factive)} \\
& \text{ for to} \\
& \text{ Q} \\
& \text{ Poss Ing} \\
& \text{ Act Nom} \\
& \text{ Der Nom}
\end{align*}
(3.175) a. Is that he'll stop at every toilet along
the way likely?
b. Is that he'll stop at every toilet along
the way ridiculous?
c. Is for him to stop at every toilet along
the way likely?
d. Is where you live known to the FBI?
e. Does his stopping at every toilet along
the way bother you?
f. Will his resisting of the police disturb
the neighbors?
g. Will his resistance of the police get him
into trouble?

IISNPC will work if it is made sensitive to the degree
of the relationship of the category label S to the node that
dominates the various complement types in (A1). In particu-
lar, the degree of violation is high for high S-ness, and
low for low S-ness.

Note that without a fuzzy node there is no possible
way the grammar, as presently conceived, can account for
the observable facts.

On a mimeographed "handout" I received from Ross,
dated March 13, 1972, the following "sloppiness" squish, to
which I have added (A2c), is given (the grammaticality judg-
ments are relative to the reading where "Bobby" and the
deleted "him" are coreferential): (A2a) ≥ (A2b) ≥ (A2c) ≥
(A2d) ≥ (A2e) ≥ (A2f).
(A2) a. Ted hates it that I pester him
   b. Ted hates it for me to pester him
   c. Ted knows why I pester him
   d. Ted hates my pestering him
   e. Ted hates my photographing of him
   f. Ted hates my contempt for him

While it is not at all clear to me how to formulate problems of sloppy identity, the data (A2) suggest that the position in the squish (A1) is a relevant factor. If the complement node is a single fuzzy node, then the degree to which it belongs to the category $S$ (or $NP$) will figure in whatever formal devices are necessary to account for speaker's intuitions on sloppy identity. If a discrete choice is forced in each case from among the possibilities $NP$, $S$, $NP[S]$, it will not only be difficult to account for (A2) and (3.175), but all the numerous other interactions of the squish (A1) will be bound by whatever decision is made, and this simply would not work.

The best "discrete" suggestion I know of, made by Emonds (1970), is that from *Poss* Ing on down, the complement node is dominated by $NP$, otherwise by $NP[S]$. This, however, seems to be controverted by the presence of a pied piping squish: In (A3) $a \leq b \leq c \leq d \leq e \leq f$.

(A3) a. This problem, that Ted solved which is impressive...
   b. This problem, for Ted to solve which would be impressive...
c. This problem, how to solve which Ted discovered yesterday...

d. This problem, Tom's solving which Ted dreamt about...

e. This problem, Tom's solving of which Ted observed...

f. This problem, Tom's solution of which Ted admired...

According to Ross's formulation of pied piping (stated in 3.3.0 as (3.199)), a dominating NP may be pied piped "as long as there are no occurrences of...the node S on the branch connecting the higher node and the specified node." This works to prevent (A3def) to be generated. This is clearly wrong. (A3a) is obviously worse than (A3c), and (A3d) is worse than (A3f). Moreover, Emonds' analysis would predict a large break between (A3c) and (A3d) that simply doesn't exist. Finally, Emonds' analysis would predict that gerunds pied pipe, whereas in general gerunds are poor pied pipers, so if his analysis is accepted, pied piping will have to be renovated.

(A4) Mary, of kissing whom Larry is not fond...

An evil act from doing which I prevented Greg...

A girl by dating whom you are sure to catch a horrible disease...

(Data from Postal (1971:134))

Note that by positing a single fuzzy node to dominate the various complement types, we can merely link pied piping to the degree of S-ness of the complement, and come out with the correct results.
These three arguments revolving about IISNPC, sloppy identity, and pied piping argue conclusively for the presence of a fuzzy node at the head of the various complement types, and against any suggestions for a discrete node such as Emonds has made.