INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Dissolving classifications: Rethinking linguistic typology

Cummings, Elisabeth Alma, Ph.D.
Rice University, 1989

Copyright ©1989 by Cummings, Elisabeth Alma. All rights reserved.
RICE UNIVERSITY

DISSOLVING CLASSIFICATIONS: RETHINKING LINGUISTIC TYPOLOGY

by

ELISABETH ALMA CUMMINGS

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

Stephen A. Tyler,
Professor of Anthropology
and Linguistics, Chair

Tullio P. Maranhão,
Associate Professor of Anthropology

E. Douglas Mitchell,
Associate Professor of Linguistics

James E. Copeland
Professor of Linguistics

Houston, Texas
May, 1989
Copyright
Elisabeth Alma Cummings
1989
ABSTRACT

DISSOLVING CLASSIFICATIONS: RETHINKING LINGUISTIC TYPOLOGY

ELISABETH A. CUMMINGS, Ph.D.

The three traditional strands of linguistics -- theoretical, typological, and genealogical -- are discussed as constructed systems of classification seeking to impose order in various ways on the world's rambunctious languages. All three strands are based on Indo-European grammatical expectations, they are enabled by literacy, and they have been empowered by a scientific mode of thought which has been dominant in the West. The postulation of "Language" as an abstract object of study is seen to emanate from an epistemology of logico-mathematics, alphabetic literacy, and the demands of a scientific methodology; note is made of the power of the Platonic metaphor. There is a growing lack of dogmatic acceptance of the three traditional linguistic approaches: the classificatory attempts to tame language are, in fact, dissolving.

A focus on typological linguistics is introduced by a tracing of typological thought from 1800 to 1963. Influences from within and from without philology and linguistics which contributed to the delineation of the subfield are commented upon. The historical and epistemological interface between theoretical, typological, and genealogical linguistics is probed.

The concentration on typological classification is continued by in-depth discussions of two languages which are of the statistically rare word order in which the object precedes the subject: Hixkaryana (Carib) OVS, and Tzotzil
(Mayan) VOS. Interpretations of sentential word order in these languages are provided from a psychological-functional approach to discourse. The order in these languages is found to constitute an anomaly from current theoretical viewpoints: they are grammatically object-subject, functionally rheme-theme, and psychologically "diffuse"-"focused". Grammatical subjects and objects are found to have limited relevance for an understanding of the discourse of these languages.

The evidence from these languages and the detailed study of the imagination which has come to dominate the study of language are meant to comprise a contribution to the dissolution of the traditional classificatory linguistic approaches. Alternatives are suggested, both implicitly and explicitly.

Unified attempts to classify the world's languages in the three traditional manners are possible; the many publications devoted to this enterprise bear witness to this fact. It is suggested that what can be principally learned from these publications is an insight into the self of Western culture, and into that mode of thought which has been dominant in the West for so long.
ACKNOWLEDGEMENTS

My years at graduate school, and the writing of this dissertation, would not have been practically possible without the gracious, patient, and loving support of my family and friends.

My first thanks will go to my mentors at Rice University, in the Departments of Anthropology and Linguistics.

It was in conversations with Dr. Philip Davis that the original idea for this dissertation was conceived. Dr. Davis’ comments, guidance, and readings of the early chapters are acknowledged with gratitude. Many hours over the last several years have been spent in Dr. Davis’ office discussing and debating bits of esoteric languages, and the nature of human intelligence. I regret that his other commitments precluded his being present at the completion of this project.

I wish to thank the members of my committee for all they have taught me, and for their unflagging support and encouragement through times both pleasant and desperate. My director, Dr. Stephen Tyler, I wish to thank for all that he has opened up for me, or, as he might say, for all that he has taught me about what I already know. Without his intellectual and practical guidance and support, my experiences in graduate school would have left much to be desired. His influence on my thought, my feeling, and my writing is pervasive.

To Dr. Tullio Maranhão I owe thanks for the constant encouragement he has given me; his intellectual influence also infiltrates my writing and
thought. Dr. Maranhao's active support for me in practical matters has more than once eased my path while in graduate school.

Dr. Douglas Mitchell's wide knowledge, his sense of humor, and his wonderful flat-footedness concerning linguistics, languages, and life, has provided me on numerous occasions with the energy and perspective needed to persevere in my studies.

Dr. James Copeland I wish to thank for his continual personal and practical support, advice, encouragement, and interest, even across departments. As well, Dr. Copeland's foundational moves at Rice over many years in the establishing and nurturing the study of language must be acknowledged with gratitude.

I next wish to thank my parents, Laurence and Jeanne Cummings. My debt to them is too great for words, in matters practical, intellectual, common-sensical, and (last but not least) financial. Their support, sense of humor, grace, and sprezzatura largely enabled the beginning, the middle, and the end of my graduate studies.

To my brother Will Cummings and to his wife Maria I also owe thanks for their unfailing encouragement, and for the loving help they have given me with my children since I began graduate school. Especially during the first two years, and at times of crises in the last three, Maria has always been there to give invaluable practical and moral support. Will has always been ready to either tender strategic support, or to enter into lively discussion of "the big picture". Without Maria’s and Will’s help, this dissertation would never have been begun, much less completed.
My best friend, Valery Roy, I wish to thank for his ability to make the dark times light for me. His steadiness and confidence has helped preserve my spirit during the last few somewhat tumultuous years.

Lastly in this list, but first in my heart, I wish to thank my children, James, Laura, and Brendan, who have been by my side throughout this adventure of graduate school and the writing of this dissertation. You were little when I began, and now are getting taller than me! You have been patient, loving, helpful, and always confident in me. You have kept me going when times were tough, and made it all worthwhile when times were good. And now I'm "Dr. Mom"! To you three, who have taught me in my heart all these years, I dedicate this work.

Houston, Texas
April 1989
TABLE OF CONTENTS

Abstract
Acknowledgements
Table of Contents

INTRODUCTION

COVERT CONSIDERATIONS .................................................. 1

CHAPTER ONE

CONSTRUCTION OF THE DISCOURSE OF LINGUISTICS

1.1. Appropriation of Languages by Means of Indo-European Grammar and
Culture ................................................................. 5
1.2. The Illusion of Discreteness: Literacy, Logic, Mathematics ............... 8
1.3. First Danger: Model Before Data .................................. 12
1.4. Second Danger: Technique Before Analysis ........................... 14
1.5. Third Danger: Fragmentation of Knowledge ........................... 16
1.6. Conclusion: Linguistics as Representation ........................... 18

CHAPTER TWO

THE CLASSIFICATION OF LANGUAGE AND LANGUAGES

2.1. Linguistic Typological Classification: 1800-1963 ....................... 20
2.1.1. Introduction: Four Considerations ................................ 20
2.1.2. Sketch of Nineteenth Century Typological Schemes ............... 24
2.1.3. Nineteenth Century Genealogical Classification: Enabling Ideology
and Effects of Triumph .............................................. 28
2.1.4. Sapir’s Voice in the Wilderness .................................. 31
2.1.5. American Structuralism, Prague School, and Joseph Greenberg ..... 38
2.2. Linguistics as Classification: Typology, Genealogy, Theory .......... 42
2.2.1. The Divorce of Talk from People ................................ 42
2.2.2. Linguistic Theory and Literacy: Construction of a Classificatory
Discourse ............................................................... 44
2.2.3. Typological and Genealogical Classification and the Scientific
Method ................................................................. 48
2.2.4. Typological Discourse: Marginality, Potentiality, and Time ....... 49
2.2.5. Platonic Thought: Guiding Metaphor in Linguistics ............... 54
2.3. The Dissolving of the Classificatory Approaches to Language(s) ....... 59
2.3.1. Questioning by Linguistic Typologists ............................ 60
2.3.2. Questioning by Linguistic Genealogists ........................... 61
2.3.3. Questioning by Linguistic Theorists ................................ 62
2.3.4. Dissolving Classifications ........................................ 64

CHAPTER THREE

OBJECT-SUBJECT LANGUAGES: A STORY OF TYPOLOGICAL RARITY
CHAPTER FOUR

HIXKARYANA (CARIB): OVS

4.1. Hixkaryana and Word Order Typology .......................... 76
4.1.2. Evidence for OVS Word Order .............................. 77
4.2.1. Hixkaryana: RHEME-initial ............................... 83
4.2.2.1. Sentence-initial Position: NEW information .......... 84
4.2.2.2. Wh- Questions and Their Answers .................... 84
4.2.2.3. Copular Wh- Questions and Their Answers .......... 87
4.2.3.1. Sentence-initial Position: Alternate Word Orders .... 89
4.2.3.2. Subject-initial ........................................ 90
4.2.3.3. Verb-initial ........................................... 91
4.2.3.4. "Discontinuous Constructions" ........................ 96
4.2.3.5. Object-initial ......................................... 100
4.2.4. TOPIC as RHEME ........................................ 103
4.3.1. Hixkaryana: IMMEDIATE-REMOTE ...................... 107
4.3.2. Direct Objects as RHEME ................................. 108
4.3.3.1. Definiteness and Indefiniteness: Identification of Particulars 108
4.3.3.2. Prefixes ................................................ 109
4.3.3.3. Particle rma 'SAME REFERENT, CONTINUITY' ........ 118
4.4. Concluding Remarks on Hixkaryana .......................... 120

CHAPTER FIVE

TZOTZIL (MAYAN): VOS

5.1. Extent of VOS Word Order in Mayan ........................... 121
5.2. Tzotzil (VOS) and Word Order Typology .................... 122
5.3. General Remarks on Tzotzil .................................. 124
5.4. Tzotzil: DIFFUSE-initial .................................... 127
5.4.1. Reports in the Literature ................................. 127
5.4.2. Expressions of Emphasis .................................. 128
5.4.3. Numerical and Quantitative Expressions: "Discontinuous Constructions" 131
5.5. Tzotzil: FOCUSED-final ...................................... 137
5.5.1. "Unmarked" and "Marked" Topic and Focus ............. 137
5.5.2. Definiteness and Sentence-final Position .............. 140
5.6. Alternate Word Orders ........................................ 143
5.6.1. Introduction .............................................. 144
5.6.2. Left-dislocation .......................................... 144
5.6.3. Particle 7a, Definiteness, and Indefiniteness .......... 148
5.7. Concluding Remarks on Tzotzil .............................. 150
5.8. "Basic Word Order", Hixkaryana, and Tzotzil ............. 150
CHAPTER SIX

TAKE IT OR LEAVE IT

6.1. A Puzzle ................................................................. 154
6.2. How Can a Language Be "Rare"? On the Concept of Markedness ... 155
6.3. On the "Unnaturalness" of Object-Subject Languages .................. 161
6.4. A Brief Deconstructive Moment ...................................... 165
6.5. The Allegory of ‘Type’ and of ‘Linguistic Classification’ ................. 168
6.6. The Anomaly of Object-Subject Languages ............................ 168
   6.6.1. Responses to an Anomaly ...................................... 168
   6.6.2. Response 1: Status Quo ....................................... 169
   6.6.3. Response 2: Compromise ..................................... 172
   6.6.4. Response 3: Revolution, or, Leave It ......................... 174
   6.6.5. Response 4: Rhetoric and Culture, or, Take It and Run .......... 175

CONCLUSION

7.1. Overt Considerations, or, Summary of Chapters ..................... 179
   7.1.1. First Chapter .................................................... 179
   7.1.2. Second Chapter .................................................. 179
   7.1.3. Third Chapter ................................................... 180
   7.1.4. Fourth Chapter .................................................. 181
   7.1.5. Fifth Chapter .................................................... 182
   7.1.6. Sixth Chapter .................................................... 183
7.2. Covert Considerations Revisited, or, Now to Begin .................. 184

BIBLIOGRAPHY .............................................................. 186
INTRODUCTION

COVERT CONSIDERATIONS

There are always more questions than answers. And, as has been remarked on by Tyler (1969b:69), if one receives an answer, it often turns out to be a function of the question asked. The very phrasing of the question is relevant in this regard.

The original motivation for my choosing to write a dissertation on linguistic typology stemmed from an awareness of the very old Greek question of "the one and the many" (this particular phrasing is due to Douglas Mitchell, personal communication). I was particularly interested in the meanderings of this question within the study of language. This is obviously a very big topic, and what is written in this dissertation is a mere entrance into further study.

The question of "the one and the many" has manifested itself in innumerable guises: essens and accidens, reality and appearance, Chomskian competence and performance, Saussurean langue and parole, the Prague school concept of "unmarked" and "marked", and so on. Nonetheless, it is a Greek question.

The particular manifestation of the question which will be primarily examined in this dissertation may be cast in terms of diversity and universality. Given the multitudinous diversity of the languages of the world, can there be any finalizing statement made about the nature of language, of languages, or of speakers? Such statements about these three matters have been attempted in various ways, within the recent study of language, and each
of them is touched on in the present work. Thus one can speak of (at least) three paths of answering in response to the inquiry posed by diversity and universality, a guise of "the one and the many".

The first path involves the possibility of making finalizing, universal, "true" statements concerning the nature of language. This has largely been the concern of linguistic theoreticians of the twentieth century.

The second path involves the possibility of making universal statements about the nature of the languages of the world. It was noted above that the manner in which a question is phrased often will determine the answer found. The seeking of finalizing, universal statements about the diverse languages of the world has most often been phrased by linguists in terms of classificatory schemes.

Historical linguists have utilized a biological and genealogical metaphor in the construction of their classificatory scheme. The implementation of these metaphors does not form a focus in the present work, but see Hoenigswald and Wiener (1987) concerning the biological metaphor, and Bloch (1983) concerning the genealogical metaphor and its early link to the doctrine of signification. The classification of historical linguists has been empowered by a heightened, culture-specific awareness of time, and by trust in the (scientific) historical-comparative method. The latter is informed by the Greek question under discussion here, particularly as it was formulated by Plato.

Theoretical linguists have, in the construction of their classificatory schemes, utilized this Platonic metaphor as well, and have postulated an abstracted "Language" which forms the basis of their scientific study. That is,
they begin with the assumption of "the one", the universally applicable statement. Linguistic theory has been empowered by Indo-European grammar, by the cognitive effects of literacy, and by the cultural dominance of technocratic science.

Not all typological linguists have sought to make such finalizing universal statements concerning the nature of the world's languages (e.g., Prague school characterology [Mathesius 1928, Trnka 1929, Vachek 1976], and Sapir [1921]). The strongest universal classificatory impulse in typological linguistics stems from Joseph Greenberg (1963), and the subsequent research on the proposals he makes in that paper. This classificatory scheme relies heavily on the Platonic metaphor, and is empowered by Indo-European grammar, and a logico-mathematical epistemology.

In short, the quest for universal statements concerning the nature of the world's languages, a quest empowered by the philosophical musings of Plato, has been conducted in linguistics by a phrasing which is notably Aristotelian: the phrasing of classification. That is to say, the enormous diversity of the world's languages have been considered by linguists through the lenses of time and fertility (genealogy), space (typology), and literate science (theory). The difference in lenses does not belie the fact that all three enterprises speak in Aristotelian, classificatory terms concerning the Greek question of "the one and the many".

The third path of linguistic investigation of diversity and universality involves the question of the possibility of making universal statements about speakers, given the diversity of the tongues they speak. This path assumes an
interrelationship between talking and thinking, or language and cognition.

"The one" along this path concerns the nature of human cognition, and attempts are made to make universal statements on this matter. It is usually assumed that diverse languages represent only "differences in expression". "The many" along this path concerns the diversity of languages, and assumes that speakers of different languages "think differently", or have experience ordered differently for them as a result of the language they speak.

It is from awareness of these very large matters, then, that my original motivation for writing a dissertation on linguistic typology stemmed. Discussion of these matters is not often overt in what follows, but I hope that the reader will find them hovering in the background.
CHAPTER ONE

CONSTRUCTION OF THE DISCOURSE OF LINGUISTICS

1.1. Appropriation of Languages by Means of Indo-European Grammar and Culture

In studying spoken language, the researcher is presented with the continuous flow of sound, and, as a literate grammarian and linguist he has learned to chop this flow into discrete, analysable entities. Davis (MS.c) suggests that anything strange and unfamiliar ("asymmetrical") is a necessary target for the focus of an intelligent being, and remains so until the unfamiliar has been accounted for somehow, and has been subsumed under what is known and familiar (that is, made "symmetrical"). Tullio Maranhão (personal communication) is speaking of the same process when he says that one must "appropriate" a phenomenon, a datum, or an occurrence in some way before it becomes knowledge, which means part of us. Linguists in the modern tradition were taught to appropriate language as an object of study by firstly writing it down (which renders the continuous auditory flow of talk into something discrete, still and quiet on a page); secondly, since the still marks on the page are (in the solitude of one's study) very susceptible to pattern-seeking analysis of various kinds, such analyses take place; finally, after the literary marks have been subjected to analysis, they are systematised or classified: patterns are described. The technique of the appropriation simultaneously creates the object and the mode of its understanding. In this way, language is appropriated by linguists and made into an object; such appropriation allows for a feeling of power over language. The ancient mystique of the spoken word
and the continuous noise of talk are both made scientific and comprehensible; they are easily stripped of their magic and of their potentiality.

The kind of classification still used by most linguists, and which may be the one thing which holds much of the discipline together, is traditional Grammar. Grammar both is a classification in itself, and it has been used as the basis for other linguistic classifications. By far most linguists have been speakers of Indo-European languages and have been formed by a Western education; they have thus been subject to the import of the Grammar and culture of the Greek and Latin tradition. Those in linguistics today who are not native Indo-European speakers join its discourse within this tradition and are therefore much influenced by it as well.

Traditional grammatical terms are still used (however awkwardly) for describing non-Indo-European languages and this poses a problem. The difficulty is acknowledged at least covertly especially by those who are dealing with those languages. One way to deal with the problem in syntax has been to concoct a description of a language's syntax in the same manner as one has been taught to "do" phonology: first, one picks a "dominant" (or rather, most frequent) particular construction and describes it; secondly, one lists occurrences which appear to be in "complementary distribution" (that is, variants of this "dominant" construction which occur in differing syntactic environments -- "allo-constructions", perhaps?); and, finally, one gets to the proverbial phonological "elsewhere", where one lists all occurrences of a particular construction which do not fit into the analysis as one has concocted it thus far (e.g., the presentation of the objective constructions in Estonian: Oinas 1966; Raun and Saareste 1965). Non-Indo-European morphology
provides similar difficulties for the analyst. For example, Uralic languages have a rich derivational morphology; the manoeuvrings necessary to "fit" Uralic morphology into the mold and classification system of Indo-European can be seen in Collinder's (somewhat redundant!) division of Uralic derivational morphology into four types: "deverbative verbs, denominative nouns, deverbative nouns and denominative verbs" (1960:220). The very division into "morphology" and "syntax" as separable arenas of analysis within a language has posed the problem of finding a finite definition of "word" and "sentence"; such a discrete definition is necessary for a structural analysis to ensue. Powell (1880:46ff.) provides an early example of a discussion whose concerns revolve partly around this question, and partly around the question of the value of Indo-European grammatical terminology and expectations when describing North American Indian languages.

Granted, it is very difficult to find a way of talking about languages which may be employed instead of the Indo-European grammatical terminology. Efforts have been made to develop some way of talking about languages which would avoid the forcedness involved both in applying Grammar to non-Indo-European languages, and in investigating the syntactico-semantics of Indo-European languages themselves. People concerned with this have proposed logico-mathematical systems and notations and diagramming techniques derived therefrom (transformational grammar, glossematics, tagmemics, stratificational grammar and subsidiaries of all of these). These kinds of approaches represent the advent of structuralism in linguistics; they have received much attention since the popular dissemination of computer technology (as linguists approach the enterprise of Artificial Intelligence).
Since both computers and these kind of linguistic approaches are based on logico-mathematical epistemology, they seem to be peculiarly suited to each other, and promise hope for those interested in attempting to make minds out of machines. Computer science coupled with logico-mathematical linguistics has been a very lucrative union. However, it has been becoming more obvious even to some of those involved in Artificial Intelligence and natural language processing that the mind and language are incompatible with the machine, and that results from this work will remain stilted and bounded (Winograd and Flores 1986). The Faustian offspring of this coupling, however wealthy, has done little but re-affirm that computers spring from logico-mathematical blood, and that the mind and language do not.

Thus, there are two major problems inherent in most modern linguistic work: first, most "theoretical" linguists have come to expect language to be fully accessible to them (from within a logico-mathematical framework), and to expect language-as-data to be compatible with a logico-mathematical system or model; second, most "practicing" linguists use, ultimately, a terminology for talking about languages which derives from a teaching technique for Indo-European languages (that is, traditional Grammar). (This same grammatical division and classification of syntax of course often also forms the basis for the logico-mathematical classifications.)

1.2. The Illusion of Discreteness: Literacy, Logic, Mathematics

Most current linguistic description, then, involves these two degrees of abstraction from the "target language" itself. Another significant degree of abstraction is (as has been mentioned) the process of writing down the
language in the first place. This process also involves pre-conceptions which color significantly the resulting de-scrip-tion or re-presentation of the language in question. The process of abstracting from the flow of sound and creating phonetic transcriptions encourages the linguist to believe that he is dealing with things, with discrete and tangible "bits and pieces". Tyler (1987:149-170) has indicated the "thing-oriented" nature of Indo-European words; it may be difficult for Indo-European speakers to conceptualize in manners that are not oriented towards discrete "things" (cf. Whorf 1956).

The mathematician goes through a similar process when he describes a physical, "real life" problem in terms of numbers. The linguist uses the letters of the alphabet, and the mathematician numbers; both (ultimately) are making models of "the real world", models that are impossible without literacy. The linguist, once he has a transcription of a language, can then perform analyses on these discrete phenomena, and the flow of sound can be put into diagrams, charts, and linguistic articles and textbooks. The process of making the sound of language visible and discrete also makes it very susceptible to the logico-mathematical kind of methodology described above. Joos, in his explicit rendering of linguistics-as-mathematics, says that of the two kinds of mathematics (continuous and discrete), linguistics is of the second sort, and should "not even make any compromise with continuity" (1950:350). How are the flow of language and the dynamics of communication to be dealt with, given this stricture? By confining one's sphere of interest to one's model:

All continuity, all possibilities of infinitesimal gradation, are shoved outside of linguistics in one direction or the other. There are in fact two such directions, in which we can and resolutely do expel continuity: semantics and phonetics. (Joos 1950:350)
Thus confined, linguistics can only be concerned with the constructions of the analyst: neither sound nor meaning are of interest!

Most linguistic theories involve and presuppose logic and mathematics; both mathematics and logic necessitate the absolute discreteness of the entities being dealt with; the illusion of discreteness is provided for the linguist by alphabetic literacy. Greenberg (1974:18) comments:

Linguistics has largely owed what success it has attained to the fact that its data, whatever their physical nature may be, have been apprehended by its practitioners in discrete rather than continuous terms. The prescientific inventions of the alphabet bears witness to the "naturalness" [sic] of this approach.

This statement from a linguist makes explicit the direct link between alphabetic literacy and the very possibility of linguistics. It also makes explicit that the understanding of "things" as discrete rather than continuous is a function of alphabetic literacy. By indicating that the "naturalness" of apprehending "things" as discrete rather than continuous is due to "the pre-scientific invention of the alphabet", Greenberg intimates that science itself is a function of alphabetic literacy. If one is aware at all of what is happening in quantum physics today, suspicions that "things" are not indeed discrete, but are continuous are causing the very foundations of Newtonian science to be questioned (Capra 1975; Jones 1982). Benjamin Lee Whorf, so often hedged by linguists, also questioned the impression of discreteness which has been so prominent in Western thought in general; he attributed this impression to the languages spoken in the West: "... English and similar tongues lead us to think of the universe as a collection of rather distinct objects and events corresponding to words" (Whorf 1956:240). He goes on to say
The Indo-European languages and many others give great prominence to a type of sentence having two parts, each part built around a class of word -- substantives and verbs -- which those languages treat differently in grammar .... this distinction is not drawn from nature; it is just a result of the fact that every tongue must have some kind of structure, and those tongues have made a go of exploiting this kind. The Greeks, especially since Aristotle, built up this contrast and made it a law of reason. Since then, the contrast has been stated in logic in many different ways: subject and predicate, actor and action, things and relations between things, objects and their attributes, quantities and operations. And, pursuant again to grammar, the notion became ingrained that one of these classes of entities cannot exist without an entity of the other class, the "thing" class, as a peg to hang on. (1956:241)

Thus Whorf would attribute the impression of discreteness to both the Indo-European languages and the cultural inheritance of the Greeks (including logic and mathematics) (again, cf. Tyler 1987:149-170), and Greenberg (by implication) points to the link between alphabetic literacy and the "naturalness" of the linguistic project, which for its very being depends upon the apprehension of discrete entities. These considerations, combined with the reification of traditional Indo-European Grammar, make linguistics as practiced a very precarious enterprise.

There are three dangers specific to the practice of linguistics when the study of language is conceived with a logico-mathematical basis. Sapir (1921:89) has said that

\[ \sim \]

destructive analysis of the familiar is the only method of approach to an understanding of fundamentally different modes of expression .... It is often precisely the familiar that a wider perspective reveals as the curiously exceptional.

The present discussion is conducted a spirit similar to that expressed in this statement, and meant as an invocation to sensitivity and to awareness of the
present situation in linguistics.

1.3. First Danger: Model Before Data

The first danger in logico-mathematical linguistics is that such theories posit the \textit{a priori} existence of a system or model before "data" may be analytically encountered. Joos (1950:350), acclaiming the "mathematical phase" in American linguistics, and dating it from Bloomfield (1926), likens mathematical and linguistic description to the use of maps:

One proceeds across the terrain and simultaneously traces a line across the map; one notes discrepancies between one's reading of the map and the sense impressions from the real world, until the discrepancies seem to form a pattern in themselves; then one corrects the map and starts all over again. All this is intuitive behavior, and logically unjustifiable. Nor does it need justification. The place for logic is inside the map, not between the map and the real world.

The question is, of course, where does the linguistic "map" come from, and furthermore, of what use is it when drawn? (This is a particularly thorny question for Joos, who disallowed the consideration of both sound and meaning from the linguistic province. What is left is then only the analyst's analysis, the "neat construction of the speculative mind" of which Sapir [1921:144] talked.) Consider Greenberg's description almost twenty-five years later of what constitutes "analytical methodology": "Ideally, by means of an \textit{a priori} logical analysis of a certain class of phenomena, a schema is developed which exhausts the logical possibilities. This schema is then applied to the empirical data" (1974:26). Greenberg goes on to say that "In fact there are probably no pure cases of the analytic approach since presumably there is always a background of some prior observations ... " (26). Thus Greenberg still posits as most desirable that the analyst approach his analysis with both logic-based
expectations and, perforce, some prior (sense-)experience with the data.

One might point out that, as Tyler (1969b:69) remarks, "... answers turn out to be a function of questions .... Whether one also wishes to believe that the data are ordered in the same manner as one's questions is largely a matter of taste." Tyler goes on to say that the "principles" of inference and induction are "a priori processes by which we categorize and generalize the partial and incomplete knowledge given in sense experience" (1969b:70). However, it is one thing to categorize and generalize in everyday life and quite another to do so within a partially or purely scientific professional endeavor (cf. Tyler 1978). Any approach to data which is lodged within or seeps out of a logico-mathematical framework works deductively rather than inductively, as Greenberg described. Compare Greenberg's description to that of Tyler (1978:274-5) as the latter describes how a typological classification is constructed:

We knew what the contrasts were going to be beforehand and proceeded deductively. It may be that we had previously done some sort of internal intercomparison which led us to think that these types would be relevant, but more than likely we had a preconceived notion of their relevance because of their role in some more general ... theory.

Comrie (1981:31) describes typological methodology in very similar terms:

In carrying out a typology of languages on some parameter, one establishes a certain number of logically possible types, and then assigns each language of the sample to one or other of these types.

Tyler makes the important remark that "Typologies are not frequently used in folk classification, and are rare outside analytic contexts" (1978:275). This latter observation attests to the fact that "analytic methodology" in Greenberg's sense emerges from an understanding based on classical or
symbolic logic (Greenberg 1974:17). It would be completely comic to claim that
humans are born with symbolic or classical logic in their brains, and that
these logics are not learned phenomena. (Cf. Luria 1976 for an investigation
into the links between literacy and logic.) If they were, of course, the coupling
of linguistics with computer science would be able to bear a sweeter child.
Thus it it is undeniable that Greenberg's "analytical methodology" (which is
an example of the whole logico-mathematical approach to language) is a
cultural construct. Both statements concerning the nature of language in
general, and the language descriptions which emerge from
logico-mathematical-scientific approaches are thus ideologically determined,
confusing and misleading. I would concur with the statement that Mallinson
and Blake (1981:14) make when discussing language descriptions which have
emerged from this kind of theory: "We hope that a just God sends the authors
of such grammars to Purgatory and sets them the task of reading their own
and each other's works".

1.4. Second Danger: Technique Before Analysis

The second danger inherent in the actual practice of logico-mathematical
linguistics is that it necessitates the mastering of a technique per se before one
can analyse a language. Now, in the learning of any technique one also,
perhaps rather insidiously, gets an idea of "how things work", or "what makes
things tick", or (perhaps more often), "how things are supposed to work". As
one first learns to drive a car, one is conscious of each movement one makes to
get the machine to drive; once the technique is mastered (once one has
"appropriated" the technique), both one's movements and the machine one is
operating become un-thought-about; however, it always somehow shocking,
and somewhat of a personal affront if one goes out one morning and the car does not start! This is because one's expectations of "how things are supposed to work", which came to be because one had mastered a technique, have been thwarted. Thus, it is impossible to learn a technique without also adding to one's understanding and thus to one's expectations. The techniques to be mastered in the linguistic approaches I have been discussing thus change the understanding of their practitioners. As students (especially those with a mathematical or computational background) learn these "linguistic" techniques, they can feel them to be somehow "natural". This feeling arises, of course, because they have been educated in a certain way: their understanding of the world, and their expectations of "how things work" have been framed by their appropriation of the techniques involved in their Euro-American education. In Davis' terminology, the "known" or "symmetrical" knowledge and feeling with which these students come into the discipline of linguistics, and upon whose "unknown" or "asymmetrical" unfamiliarities their intelligence comes to focus is constituted by their prior education. This education has been had within cultures whose dominant ideology has been scientific, so that it is easy for their intelligence to interpret linguistics simply as a slice of the same pie: its taste is familiar.

Consider the following statement:

There would seem to be a certain 'naturalness' about the assumption that the set of domain members should be individuals while the range set should consist of properties including, on occasion, numbers, i.e., metrical properties. (Greenberg 1974:21)

Now, obviously there is nothing "natural" about this assumption unless one has appropriated set theory and now has one's expectations guided by it.
Greenberg's statement is an extremely explicit example of the kind of understanding of students just referred to. If the student has concentrated on mathematics, science, and computer science, he will feel an immediate affinity to the linguistic theories and modelizations; but even if he does not have such a background, chances are that he will accept the general premises of the theories because they are but explicit structures mirroring the way he has been taught that "things are". An analogy can be made from architecture. Consider the Pompidou Centre in Paris: the externalization of the structural components strikes one as a slightly odd, but since it has uncovered the structure which one always supposed was there, it is not totally bizarre and alien. Because students have been through the Euro-American educational system, they have inherited a certain expectation of "how things are", and thus are not often critical of the logico-mathematical linguistic theories; they may at first feel vaguely uncomfortable with what has been called "the mathesis of language" (Tyler 1978:xi), but after a while the feeling goes away, and everything seems to be alright. Things all fit together somehow.

1.5. Third Danger: Fragmentation of Knowledge

The third danger which is inherent to the practice of the logico-mathematical linguistic theories is one which is not confined to linguistics; it, along with some of the other considerations just mentioned, riddles our Western culture. Lyotard (1984) has referred to this danger as "the fragmentation of knowledge". By this is not implied that "there was some commonality of 'knowledge' in an earlier and perhaps romanticized version of Western civilization" (Philip Davis, personal communication). Rather, I believe, it is a simple synchronic statement of where we find ourselves today.
Though I would not invoke a strict developmental sequence leading to our present day, I would suggest that there have been some surging impulses which have helped to bring us to where we find ourselves. A beginning list of these might include the following: a progressive division of labor; very rapid advances in technology and industrialization, (including the invention of the printing press, the dissemination of literacy, and the advent and dissemination of the electronic communicative media); the increasing concept of the self, from the man-centering of the Renaissance to Freud's ego; the secularization which accompanied this increase of "self"-awareness; and finally, the dominant epistemological framework bequeathed by a resurrected Aristotle who was then appropriated and disseminated by Descartes. Thus the fragmentation of knowledge does not only refer to the specialization of "information" incumbent on any professional today. Specialization in "careers" in the extent to which it is extolled today is fragmenting. An obvious example is that even specialists in the same field cannot communicate with each other without learning one another's jargon. Those in other disciplines, and the general populace, are completely excluded from communication with the specialist. Another example is the extraordinarily isolating and compartmentalized daily life in North American cities -- this kind of daily life and the results of the idealization of career specialization are symptomatic of the "fragmentation" of which Lyotard speaks. And they are but another instance of the playing up of the "discrete" and the downplay of the awareness of continuity. Thus linguistics as a scientific endeavor is but a a symptom of a wider dis-ease in our culture.

1.6. Conclusion: Linguistics as Representation
There are two ways in which the logico-mathematical linguist deals with the object of his study, language. One of these is static and structural; the other involves "relations" and "functions" (in the mathematical sense of these terms) between elements, or discrete entities. Often these two ways are combined, but not always. Both are impossible to conceptualize or to practice without prior formalization or modelization; they are dependent on logico-mathematics, visualization and literacy. They are rich examples of the representational mode of thought. If one "re-presents" something, one presents it again, shows it, or displays it a second time. By definition, the "it" is not the same thing as the representation. Representational thought relies on the insistence that there is something out there which is "other than me"; that something, the "other", one may observe, and one may reach, principally by an effort of reason; it is a mode of thought which insists that one can capture what one has observed, and present it again, make a picture of it, or talk definitively about it. If this way of thinking is applied to language, certain assumptions are immediate: language is outside oneself; language is observable; language is able to be caught and represented: one can make a picture of language, and talk definitively about it.

Representational (or logico-mathematical) linguistic theories have, in practice, become less concerned with the actual languages of the world and more concerned with the attempt to improve the representation of "Language", the latter having been apprehended as a metaphysical object. As Mallinson and Blake (1981:27) have said,

Above all we should be wary of imposing on the data we discover a preconceived notion of the most suitable theoretical framework .... Unfortunately, some linguists have attempted to fit the data to the theory rather than
the other way round.

The imprecision of the modelization can be admitted by logico-mathematical theoreticians; they fuss about how to make the modelization more precise while preserving simplicity, elegance and economy. Their discourse is of the sort described by Tyler (1987:201):

As science increasingly defined itself as the mode of discourse that had its own discourse as its object, every move to perfect that discourse and fill every gap of proof revealed ever-new imperfections. Every self-perfecting, self-corrective move created local orders that spawned new imperfections requiring new corrections. Instead of a coherent system of knowledge, science created a welter of local orders unrelated to one another and beyond the control of anyone. Scientific knowledge was systematized only by the unity of a rational method that produced greater and greater irrationality.

It is a self-reflexive discourse from which, like a dying star, no light can escape: it is folding in on itself, collapsing from its own weight.
CHAPTER TWO

THE CLASSIFICATION OF LANGUAGE AND LANGUAGES

2.1. Linguistic Typological Classification: 1800-1963

2.1.1. Introduction: Four Considerations

Joseph Greenberg is one who, amidst the plethora of representational, theoretical linguistics of the last half-century or more, makes an attempt to "get back to the data", to focus attention once again on the actual languages of the world. One might point out that the intent behind theoretical and typological linguistics differs, though both purport to deal with language(s). Given the enormous diversity provided by the languages of the world, theoretical linguists seek to create a scientific theory and modelization which will "account for the data", and pay attention to the criteria of exactness (including predictive power), generality, and simplicity (Davis 1973:1-12); thus, as I have already indicated, the discourse which results is self-reflexive and self-perpetuating. Typological linguistics, on the other hand, given the same enormous diversity found in languages, has remained comparatively much "closer to the data", to the phenomena of languages. Rather than postulating and concentrating on structural abstractions (langue, "competence", habit, relational networks, over-lying structures or underlying forms), this approach has been content to a much large extent with parole, with "performance" (I use these terms for convenience; I consider the distinction itself to be questionable, and at best unimportant). Despite the almost overwhelming diversity presented by a survey of even the documented languages of the world, this approach groups languages together according to
formal similarities or formal patterns observable in them, and thought to be comparable. The reasoning process of theoretical linguistics, one might say, is deductive, while, initially at least, that of typological linguistics is inductive.

There are essential differences between the typological classification of languages prior to Greenberg’s important 1963 paper (1963b), and the kind of classification to which he brought attention. There are four considerations which will be interwoven in the following discussion of the history of typological classification: the unit of classification, the method and reasoning process employed in the construction of classifications, the ideological climate, and the research climate.

The first consideration involves the question of what is to be classified. It has been sometimes noted that "... the history of typological investigation has been in large measure ... the quest to determine which aspects of language will bear the most fruit on a cross-linguistic basis" (Mallinson and Blake 1981:19). Early nineteenth century language scholars based their typological classifications on the structure of the word (morphology). With the development of Prague school phonology in the late nineteenth and early twentieth centuries, the unit of classification became much smaller than the word: phonetic features. When linguists became ultra-aware of syntax in the late 1950's, the unit of classification became the sentence, but often the methods employed were those of Prague school phonology, both in syntax and in semantics. Sapir (1921) had the most holistic conception of language, one which gives consideration to cognition, semantics, and rhetoric. His morphological classification alone goes further than considering only the mechanical, formal structure of the word.
The second consideration involves the mechanics of the construction of the classification. I will invoke the famous logical dichotomy of inductive and deductive reasoning here because (though I consider it to be an empty formalism in everyday life), within research, one process has been more acceptable than others at different times. This changing popularity has influenced the idea of what will or will not constitute a "proper" classification. The reasoning process which is acceptable at a given time informs the method with which "data" of any kind are approached. In the study of language in the last two centuries, there has been a steady growth of faith in the fertility of deductive reasoning, almost to the withering out of any inductive reasoning at all. Only in recent years has this freakish un-balance begun to be wondered at. Although the dominance in the United States between 1933 and 1957 of the behaviorist-descriptivist approach is apparently inductive, this is so only if induction is equated with empiricism. The Bloomfieldian militant mechanism does not preclude its primarily deductive mode of reasoning.

The third consideration involves the philosophical, literary, economical, political, and ontological (religion, magic, and science) moods within which research is conducted. All these have an effect upon both the interests and the expectations of a researcher, and upon the acceptability or non-acceptability of his research. To some extent the interests of the researcher are limited by this ideological climate, by the "current discourse" on a particular topic. Gadamer (1983:36) has pointed out that

It is precisely our experience of history that we are so completely within it that we can in a certain sense always say, We don't know what is happening to us. History consists precisely of the fact that we do not realize what is happening to us and that nevertheless we are involved in this play ...
Ideological expectations thus characterize the outcome of a researcher's work; language classification has not been immune to these expectations, since they in fact determine to some extent what will be the "relevant" questions to be posed.

The fourth consideration is the fact that linguistics is a new academic discipline within the university. Research in any field is a dialogical activity: talking with and listening to one's colleagues (in person or in print) is an integral part of research life, and of the process of research. This talk and listening necessarily influences any research that is done. Within the university, the study of language has become increasingly detached from all other fields, an obvious accompaniment to the much-heralded establishment of "linguistics" as a discipline with its own academic departments. This change in research climate has had a tremendous effect on the research of those interested in language and languages. For instance, Hymes (1974b:10) notes that

it is no mere coincidence that the dominance of the 'neo-Bloomfieldian' approach came with the emergence of linguistics as a distinct academic profession, separate from philology, language departments, and anthropology departments, in the United States.

Consider, for a moment, the implications of the change in how people interested in the study of language(s) are commonly referred to: those in the nineteenth century are referred to as "language scholars", or "philologists"; those in the twentieth, as "linguists". The kinds of classificatory schemes proposed over the last two hundred years reflect this change.

With these four considerations in mind, what follows is a selective narration of the research in the typological classification of languages in the
last two centuries. I mention these considerations, and yet must embark on such a selective narration; it is beyond the scope of the present work to further delve into the history of linguistics. I refer the reader to Hymes (1974a), to Aarsleff (1982), to Hoenigswald (1979), to Hoenigswald and Weiner (1987), to Lounsbury (1968), to Voegelin and Voegelin (1963), and to Sebeok (1966) for more detail. However, I do concur with Hymes (1974b:2), when he says that

the complex history of linguistics, closely considered, resolves into an overlapping series of local scenes, specific 'structures of feeling' ... approachable through biographies ... and lesser writings, more than through isolated classics .... Isolated works may suffice, if our interest in the past is only to praise or condemn. If our interest is to know what happened, we must enter into the contexts of personal and institutional origins. Our history must become a history, not only of great men, but also of circles, and not only of circles, but also of institutions, governments, rulers, wars and the ways in which these have shaped the renewed origins of linguistics in successive generations.

And now, to my narration; I rely heavily on Jespersen's 1920 essay for much of the nineteenth century German research.

2.1.2. Sketch of Nineteenth Century Typological Schemes

In 1808, Friedrich von Schlegel, one of "the chiefs of the Romantic movement", on the basis of a "rather superficial knowledge of a small part of the languages of the earth", proposes that all languages may be divided into two groups, "organic" and "mechanic". The organic languages are comprised of Sanskrit and its cogeners; they are characterized by "organic" growth of their roots, since they are capable of inner change ("flexion"). The mechanic group is comprised of all other languages; they are characterized by affixation. Jespersen makes two comments on this: 1)von Schlegel's arguments for including Semitic among the affixing languages are weak, and Jespersen feels
that von Schlegel is rather inclined to say that Semitic involves "real flexion"; 2) there is a germ of a tripartition in the classification, since on the lowest level of the second class von Schlegel places Chinese, in which "particles denoting secondary sense modifications consist in monosyllables that are completely independent of the word itself" (Jespersen 1920:694).

Friedrich's brother, August W. von Schlegel, in 1818 "gives voice to" such a tripartite division, recognizing that Chinese works in an entirely different way from either affixation or inflection. Thus the first class he proposes is comprised of languages with "no grammatical structure", like Chinese. The second class is comprised of languages with affixation. The third class (which is also the "highest" class, and whose members alone are "organic") is comprised of languages with inflexions. This class he further subdivides into synthetic (e.g. Latin), and analytic (e.g. Romance languages); the latter are languages created in modern times, while the origin of the former "se perd dans la nuit des temps" (Jespersen 1920:695).

Franz Bopp's contribution shifts the emphasis somewhat; he still uses a unit for his classification which is word-internal, but rather than concentrating on the additions to a root (in Sapir's terms, the "technicalities" of "relational concepts"), he uses the sounds of the root itself as his basis (a phonetic approach perhaps pre-figuring that of the Prague school); Jespersen calls Bopp's classification "root theory", and refers to a letter in which Bopp claims that his theory makes von Schlegel's work "fall completely to the ground". Bopp's fully developed classification appears in 1857; he also ends up with three divisions, even though the basis for the classification is somewhat different from the von Schlegels'. The first class is comprised of languages
without composition, such as Chinese. The second group is comprised of languages with monosyllabic roots, which are capable of composition, and obtain their grammar almost exclusively in this way; included in this group are the Indo-European languages, plus all languages not in the first or third classes (!). The third class is comprised of languages with disyllabic roots and three necessary consonants "as exclusive bearers of the signification of the word"; they involve composition and inner root modification; this class is comprised of the Semitic languages only (Jespersen 1920:695-696).

Jespersen notes that neither with Bopp nor with the brothers Schlegel does the idea appear that these morphological divisions of the languages of the world are stages in historical development. This idea, he says, first finds articulation with Wilhelm von Humboldt.

Jespersen refers to Humboldt's 1822 paper, in which Humboldt suggests that between the two typological poles of Sanskrit and Chinese lie all the other languages of the world, and that there has been a four-stage progression in the development of languages. In the first stage, language has "no grammatical form", and "denotes only objects, leaving it to the hearer to understand or guess at ... their connexion"; Chinese is the prime example. In the second stage, "word order becomes fixed", and "some words lose their independent use", so that "grammatical relations [are] denoted through word order and through words vacillating between material and formal significations"; these Humboldt terms incorporating (e.g., Mixteca). In the third stage, these "vacillating [words] ... become affixes, but the connexion is not yet firm, the joints are still visible, the result being an aggregate, but not yet a unit"; this type of language, the agglutinative, has "something analogous to form, but not
real form" (Jespersen 1920:696-697). Horne (1966:14) says that Humboldt considers these languages to have "intended but incomplete flexion, a more or less mechanical affixing", but that they do not show "a truly organic development". The fourth stage of language development, typified by Sanskrit, shows "true organic development" (Horne 1966:14), and is characterized by "real form": "the word is one [sic], only modified in its grammatical relations through the flexional sound; each word belongs to one definite part of speech and the words used to denote form have no longer any disturbing material signification, but are pure expressions of relation" (Jespersen 1920:697).

Humboldt’s typological classification, then, is placed within a developmental framework; this type of evolutionary approach, along with the mechanical and organic metaphors he uses, have fallen into apparent disrepute today. However, the observations which he makes about different language types are interesting for several reasons. He is the first to begin to look further than the unit of the word as a basis upon which to make classificatory statements. He also does not simply look at morphology as a form, but also as a bearer of meaning (in fact, his characterization of these four stages and types can be seen as a precursor to Sapir’s classification). Furthermore, Humboldt makes explicit that he is not claiming that there are clearcut boundaries between the four types he mentions; he says that "all languages contain one or more of these forms", and in fact concludes that "the structural diversities of human languages are so great that they make one despair of a fully comprehensive classification" (Jespersen 1920:698). Ramat (1987:194) notes that although Humboldt, like Bopp and Friedrich von Schlegel, regarded Sanskrit as "the optimal language form", he did not "see an
inevitable, continuous 'decadence' in the development of languages", as his contemporary Indo-Europeanists did. Rather, "his idealism led him to see the effect (and, at the same time, the cause) of spiritual progress in the changes in historical languages, the achievement of their implicit potential ... and hence a ... completion".

2.1.3. Nineteenth Century Genealogical Classification: Enabling Ideology and Effects of Triumph

Various modifications on this three or four part morphological typology were proposed by Bopp, Pott, Schleicher, Steinhthal, Muller, Misteli, Fortunatov, Finck, and Marr (Horne 1966). But no really different proposals were made; typological classification had really lost interest for language scholars as another way of classifying the world's languages caught their attention: genealogical classification. With the discovery of the concept of phonetic law early in the nineteenth century, culminating in the neogrammarian triumphs of the 1870's, genealogical classification became a fruitful and tangible area of research. The study of language became a scientific project, and proudly joined the "social sciences". The Romantic, naturalistic climate beamed on the idea of a search for the origins of languages, and their posited evolution in time; later in the century, the positivistic climate beamed on the "results" (etymological and grammatical reconstructions) achieved by means of the fixed rules of the comparative method. Greenberg (1974:58-59) claims that, except for Schleicher, all who worked in language typology during this time assumed that "languages which were classified together typologically could ultimately be shown to be related genetically", and that when the neogrammarian historical-comparative
"results" proved this assumption no longer tenable, typological classification ceased to be pursued. He says that Pedersen "sees typology as an attempt to classify languages on the basis of incorrect premises". I would tentatively suggest that this is an attempt by Greenberg to find a logical reason (long after the fact) for the lack of interest in typology at this time; he is also trying to support his statement that at the time of his writing (1974), there existed a "logical independence of genetical and typological definition" (1974:58).

Ramat (1987:194-195), trying to account for the same lack of interest in typological classification after Humboldt, mentions the "clearly positivistic leanings" of comparative linguistics, and infers that it was too "intent ... on reconstructing the phonetic and grammatical history of individual languages and protolanguages" to have time to deal with "the problems of typology". He goes on to say that

The memory of judgements about this or that language which had been more or less gratuitously formulated on the basis of purely impressionistic and subjective data ... certainly harmed typology.

He notes that "the entire French sociological school maintained a skeptical attitude even towards typology in the strict sense", and quotes Meillet's comment that "The only language classification which is of any value is a genealogical classification". Ramat's accounting for typological classification being uninteresting to the language scholars of this time is thus different from that of Greenberg. Greenberg is still working from within a positivistic framework, and so seeks a logical explanation from within the discipline; he indicates a logical contradiction in terms, and this contradiction explains for him why typological classification fell by the wayside, and it also legitimizes his own approach to typology, as he does not make the "illogical" claim
Ramot, however, speaks of influences from outside the discipline of linguistics and the research on language which he feels may have contributed to the disinterest in typological classification which occurred in the latter part of the last century, and which continued into the twentieth:

despite the common comparative base and the greater antiquity of language typology as compared with historical and comparative linguistics, typology remained on the sidelines of mainstream linguistics. The philosophical speculation which lies behind the interest in language diversity ... was substantially extraneous to the scientific positivism of the second half of the 19th century and even at the beginning of this century, when the dominant scientific approach in linguistics was based on neogrammarian historical reconstruction. (Ramat 1987:195)

I would like to make more explicit the point that a spirit of Romanticism given rigor by the logic of positivism contributed largely to the success of genealogical classification; tangible "results" were able to be had by any who would discipline themselves to the scientific (comparative) method, and there was never a shortage of exotic data, as European colonialism was flourishing. The successes of genealogical classification, taking place in such a climate precluded "serious" interest in typological classification for several reasons. The Romantic spirit yearned for a past Golden Age, and delighted in a search for origins. The positivistic spirit and the scientific method required the manipulation of facts, and speculation beyond the data, or contemplation, or play, became superfluous and foolish. The study of language became "rigorous", and more and more, language came to be considered as an autonomous object, as some thing which had an analysable existence of its own. The methodological fetish with rule-writing and prediction which still thrives
in most theoretical linguistic quarters today crept into the study of language at this time, with the initial successes of genealogical classification. The problem with typology as classification was in the nature of the enterprise itself, as Humboldt and others had formulated it: one can descry typological tendencies in morphology, but not absolutes; the divisions into types always had fuzzy edges, and fuzzy edges were anathema to a scientific, positivistic study of language.

2.1.4. Sapir's Voice in the Wilderness

It is Edward Sapir (1921) who revives interest in typological classification, and who gives it its most significant articulation. Greenberg (1974:41) calls Sapir "very much a transitional figure"; I would suggest that a more accurate characterization of Sapir would be that of a voice crying out in the wilderness. Sapir was not working exclusively within the scientific tradition of the genealogical classification of languages; he was not an "armchair linguist", but did extensive fieldwork with Amerindian languages; his book Language appeared in 1921, almost a hundred years after Humboldt's original formulations, and he worked in America, not in Germany. All these things permit his typological work and his general study of language to appear highly original. Ramat (1987:195) says that "Sapir felt the inadequacy of classifications based on the analysis of forms and meanings which easily strayed into psychologism ... [like Humboldt, and] thus suggested a typological classification based on formal criteria", and that the latter were suggested to him by his study of Amerindian languages. It is difficult to see how Sapir's classification can be considered to be based solely on considerations of linguistic form; its very complexity derives from the fact that Sapir was
considering not only form but meaning and function as well. Furthermore, Sapir is not so dry as to take language as a scientific object; his "psychologisms" are frequent and suggestive, and though they are of a very different cast than those of Humboldt, one can delight in their presence, surely.

One of Humboldt's impulses for his interest in typological classification was his belief that language was a way in which the spirit of a people manifested itself (Ramat 1987:199). The nature of his "psychologisms" take on a cast appropriate to this, and they are phrased in a manner appropriate to his time. Sapir, as a "practicing" "ethnolinguist" in early twentieth century America, casts his "psychologisms" in a different way (for instance, he does not suggest an evolutionary sequence in language types). His thrilled curiosity and suggestive excitement with the diversity in languages largely motivate him, and his "psychologisms" tend to be phrased in psychological terms rather than in the almost visionary historical-philosophical terms of Humboldt. But for all that, there are similarities in their approaches.

Sapir is not overwhelmed by neogrammarians method; his interest in Language (1921) is in language (as a human phenomenon) and languages (as diverse ways of expressing similar things). His interest in typology is not to codify but to classify, and his reasoning process is primarily inductive rather than deductive. The earlier typological work of the nineteenth century was also primarily inductive (Greenberg calls it "intuitive" [1974:25]). Its conclusions necessarily differ from those of Sapir simply because less was known about the world's languages: Sapir's "data-base" was obviously extensively larger. The scientific approach involved in the comparative method is, by contrast, primarily deductive (Greenberg's "analytic" [1974:25]),
as I have mentioned already, and this is why both Sapir and the early
nineteenth century typologists are regarded with such condescension by most
"scientists" of language. Thus not only the ideology, and the method, but very
reasoning process which first became dominant in late nineteenth century
Germany was incompatible with typological classification (since the latter
remains empty statistics when approached deductively).

It is well-recognized that Sapir provides an innovative approach to
Horne stresses the fact that Sapir's proposed classification "offers a wide range
of possibilities", noting that "even the most elaborate of previous typologies,
such as those of Finck, Steinthal, and Misteli offered only eight classes". He
calculates the interplay of Sapir's three distinctions (conceptual type,
technique, and degree of synthesis) and reports that Sapir proposes then "the
possibility of 2640 to 2870 language types, even without resorting to Sapir's
qualifying notations of 'mildly', 'weakly', or 'strongly' before the degree of
synthesis"; if these are calculated, he says, "there is the numerical possibility
of assigning a separate and distinct language type to every known language"
(Horne 1966:32). The fact that Horne actually made calculations of what
Sapir (1921:146) called "a few schematic indications" shows him to be one who
is uncomfortable with fuzzy edges, one who is informed by the positivistic,
logico-mathematical approach to language which had become fully dominant
by the 1960's. Horne's expectations of what constitutes a "proper"
classification are deaf to Sapir's explicit warning against such expectations: "
... classifications, neat constructions of the speculative mind, are slippery
things" (1921:144). Horne's perspective is highly different from that of Sapir:
the former checks the calculations because he has been educated to expect that 
each thing has its proper place, and that a classification should reflect clearcut 
boundaries; the latter proposes his classification not as an end in itself, a 
reification, but as a useful tool, as a way of talking about the diversities of 
language:

    It is of less importance to put each language in a neat 
pigeon-hole than to have evolved a flexible method which 
enables us to place it, from two or three independent 
standpoints, relatively to another language. (Sapir 1921: 
140)

Greenberg (1974:41) refers to Sapir as a transitional figure because, while 
Sapir "firmly rejects both the valuative and evolutionary aspects" of his 
precursors, he still "evidently feels that the structure of the word is the central 
consideration characterizing languages and shows ... sympathy with the notion 
of languages as an expression of the spirit of a people." I have already 
indicated that both Sapir and Humboldt do indulge in such "psychologisms". 
However, again it is difficult to see how Sapir can be interpreted as giving to 
"the structure of the word" (that is, to formal morphological criteria) the 
centrality in language that Greenberg suggests. Sapir's initial considerations 
and subsequent groupings into four types could simply not have been done 
without consideration of morphological meaning and function. Sapir explicitly 
states the following:

    This conceptual classification of languages, I must repeat, 
does not attempt to take account of the technical externals 
of language. It answers, in effect, two fundamental 
questions concerning the translation of concepts into 
linguistic symbols. Does the language, in the first place, 
keep its radical concepts pure or does it build up its 
concrete ideas by an aggregation of inseparable elements 
(types A and C versus types B and D)? And, in the second 
place, does it keep the basic relational concepts, such as 
are absolutely unavoidable in the ordering of a
proposition, free of an admixture of the concrete or not (types A and B versus types C and D)? (Sapir 1921:138)

This is the heart of Sapir's typological groupings; it is only as an "amplification" that the "sub-divisions" into agglutinative, fusional and symbolic "sub-types" are made. Thus, consideration of formal word-structure, involving "the prevailing method of modification of the radical element" (1921:138-139) appears only as a finer tuning of the original "conceptual" divisions.

Ramat (1987:20) refers to Lehfeldt and Altmann (1975:73ff.) who consider four "stages in the development of typology" and who "identify four conceptions of 'type'":

a) the classificatory definition of the Schlegel brothers who group classes according to the absence or presence of a single specific feature ... ;

b) the combinatorial definition (Sapir) which defines classes by means of a combination of several features;

c) the definition which takes type to be a scale between two extremes along which languages are to be found according to the degree of development of the characteristic examined (Humboldt): this approach leads to an understanding that classification pure and simple cannot be the be-all and end-all of typology;

d) the final definition considers type as an ideal model and ordering principle ... which, according to Lehfeldt and Altmann, can be used efficiently only when concepts of measurement are introduced ....

One might say that Horne is calculating in order to see if Sapir's scheme conforms to the positivistic idea of what a classification should do: have clear-cut "results". Greenberg's comments on Sapir are directed towards consideration of the unit of classification. Lehfeldt and Altmann are considering differences in how these typological classifications were formulated:
the Schlegels’ on the basis of the presence or absence of features, Sapir’s on the basis of intersecting features, Humboldt’s on the basis of a developmental scale, and the logico-mathematical one on the basis of a preconceived model and ordering principle called "type".

If one looks only at the chart which Sapir gives (1921:142-143), and the explanation of its terms, one is confined to considerations of the kind just discussed. However, Sapir’s table is just one part of his conception of language. It needs to be explicitly noted that Sapir not only does not base his morphological classification on the consideration of formal elements, but that he does not consider the word in isolation at all. He repeatedly emphasizes the fundamental importance of both sequence (today called "syntax") and stress (today part of "suprasegmental phonology", formerly part of "prosody"). He emphasizes these both "within" the word ("The elements of the word are related to each other in a specific way and follow each other in a rigorously determined sequence" [1921:110]), and "outside" the word:

What ... are the fundamental methods of relating word to word and element to element ... of passing from the isolated notions symbolized by each word and by each element to the unified proposition that corresponds to a thought? .... The most fundamental and the most powerful of all relating methods is the method of order. (1921:110-111)

And again:

It is somewhat venturesome and yet not an altogether unreasonable speculation that sees in word order and stress the primary methods for the expression of all syntactic relations and looks upon the present relational value of specific words and elements as but a secondary condition due to a transfer of values. (1921:113)

In his discussion of grammatical "concord", which is, of course, typical of the traditional inflectional languages, he says:
Psychologically the method of sequence and accent lie at the opposite pole to that of concord. Where they are all for implication, for subtlety of feeling, concord is impatient of the least ambiguity but must have its well-certified tags at every turn. Concord tends to dispense with order .... [however still] at some point or other order asserts itself in every language as the most fundamental of relating principles. (1921:116)

These quotations should make it clear that Sapir certainly does not consider the formal structure of the word as "central" to a conception of language. Perhaps, in fact, he has more in common with Ramat (1987:15), who, in his invocation of a "polythetic typology", says that

a linguistic system is not so much a system of 'data' but rather a system of strategies and techniques which resolve the problems arising from the need to express oneself.

Morphological meaning (Sapir's "conceptual", today's "referential" plus Sapir's "relational", today's "functional"), "technical features" (1921:145) (Sapir's criteria of synthesis), and rhetorical considerations (Sapir's "order" and "stress") all contribute to Sapir's conception of language. Furthermore, I would suggest that the interest today in "pragmatics", and "discourse", and "strategies of communication" (Ramat 1987; Seiler 1978; Hopper and Thompson 1980 and 1984; Davis MSS. a, b, c) are all concentrating on the rhetorical aspect of language. It seems that this current interest is directly related to Sapir's suggestion that languages work within "fundamental form intuitions", and that there are "great underlying ground-plans", of which linguistic data are merely "symptoms" (Sapir 1921:144). Ramat (1987:12) describes the suggestion of Birnbaum (1970a:25ff.) that "deep structure" is stratified into at least three levels (Ramat also notes that Dezso 1982:16 uses a similar approach):

a first level which is not very deep ('shallow deep
structure') and which characterizes a specific language; a second level, possibly subdivided into several layers, of a deeper and more general nature, characterized by a series of properties shared by language groups, which goes by the name of 'typological deep structure'; and, finally, a third level, the 'profound structure', where the deepest level and greatest degree of generality are achieved and where the properties of human languages as such are to be found.

Though he does not employ the generative terminology, nor the layer-cake model, Davis (MSS. a, b, c) has made suggestions as to the nature of this "profound structure"; he refers to it as the way human intelligence plays itself out in languages.

2.1.5. American Structuralism, Prague School, and Joseph Greenberg

Between the publication of Sapir's *Language* in 1921 and the conference on language universals in 1961 at Dobbs Ferry, at which Greenberg delivered his paper on word order typology (published in 1963), there was again a period during which not much interest was shown by American linguists in the typological classification of languages. Linguists were by this time fully established as social scientists, and language had become fully objectified; independent Linguistics Departments had been instituted in American universities. Bloomfieldianism, structuralism (or, "what linguists made of Saussure"), functionalism, and later generative grammar (full-blown positivistic logico-mathematics) were dominant movements. Sapir's suggestive work, his voice in the wilderness, was not listened to in linguistics. For instance, Joos (1958:25) (commenting on Sapir 1925) speaks of the "brilliance" and "genius" of the person of Sapir, but also mentions "the essential irresponsibility of what has been called Sapir's 'method'", and points out that Sapir "said many things which are essentially uncheckable ('invulnerable') and
thus not science". As Hymes (1974b:11) has noted,

In the time-honored manner of economically and politically successful approaches, members of the inner paradigmatic community ... [can] afford to mock ... [those not accepting the authority of a dominant approach], rather than have to take them seriously.

In Europe, typological research was not as obviated. The Prague circle's emphasis on synchronic ("static") cross-linguistic comparison (as opposed to historical, "dynamic" comparison) included what Mathesius (1928) called "linguistic characterology". Mathesius differentiated characterology from typology, saying that "All attempts at a systematic linguistic typology are, at the present state of our knowledge, premature and lead therefore to unnecessary complications of problems only" (1928:59). He describes characterology in relation to both descriptive linguistics and historical linguistics. It is the task of "descriptive grammar", he says, "to give a complete inventory of all formal and functional elements existing in a given language at a given stage of its development" (1928:59). However, linguistic characterology

deals only with the important and fundamental features of a given language at a given point of time, analyses them on the basis of general linguistics, and tries to ascertain relations between them;

he continues:

Comparisons of languages of different types without any regard to their genetic relations ... considerably furthers the right understanding of the real nature and meaning of the analysed linguistic facts. (Mathesius 1928:59)

The basis he proposes for cross-linguistic comparison is "common grammatical functions". With regards to historical linguistics, Mathesius says that linguistic characterology has the "ability to discover new problems for
historical investigation or to show new ways for the solution of problems already under discussion" (Mathesius 1928:60).

As Hymes (1974b:11) says, during this period, "structuralists everywhere concentrated on phonology"; in Trubetzkoy's phonological classification Grundzüge der Phonologie, published after his death in 1938 (Trubetzkoy 1939), one can see a combination of the phonological interest with the influence of linguistic characterology. Hockett's Manual of Phonology (1955) bears a great debt to the Prague school, and to Trubetzkoy's phonological classification. Martinet (a young member of the Prague school) makes proposals for dealing typologically with the lexicon, morphology and syntax (Martinet 1962:87-102), discusses Trubetzkoy's "pioneering efforts" in phonology, and refers to his own attempts at a typology of prosodical features (Martinet 1962:84).

Greenberg, however, returned to Sapir, for both his 1954 paper and his 1963 paper seem to be elaborations on Sapirian suggestions (often with the methods of the Prague school). Greenberg (1954) provides a method for calculating the degree of morphological synthesis in a language; it was an attempt, Greenberg later says (1974:48) "to provide ... rigor and exactitude" to Sapir's version of morphological typology using "the analytical tools of contemporary American structuralism". It is curious that Greenberg chooses Sapir's "synthesis" as the part of Sapir's scheme upon which to elaborate, as it was the one part which Sapir did not consider as fundamental. Greenberg later (1974:48) partially admits this:

the particular aspect of the typology on which Sapir placed the greatest value, the expression of derivational and mixed relational concepts, was lacking in my
treatment, because it involved an approach from the
direction of semantics.

He attributes his "lack of treatment" to the fact that he was working within
the influence of American structuralism.

In his 1963 paper, which has been quite influential in theoretical
linguistics and in classificatory linguistics (both typological and genealogical),
Greenberg elaborates on another of Sapir's suggestions: that sequence is
"fundamental" and "powerful" in language. However Greenberg's work on
word order typology is very limited for several reasons.

The first is that he takes as the unit of classification the "unmarked",
"natural" order of elements within the simple sentence. The second is that,
instead of taking "words" out of context, his method takes "sentences" out of
their communicative context. The third reason is that the lense through which
Greenberg looks at sentences is that of traditional Indo-European Grammar.
His classification and "universals" are dependent upon the assumption that
the Subjects, Verbs and Objects, and the various parts of speech and
constructions of Grammar are "universal", that they are to be found in every
language in the world. (Sapir, by contrast, seems to go to some length to avoid
concentrating on these appellations from Grammar.) In recent years, many
studies have been made which indicate the severe problems associated with an
assumption of the universal applicability of these terms and concepts from the
Indo-European grammatical tradition (Davis, MSS. a, b, c; Schachter 1976; Li
and Thompson 1976; Comrie 1981). The fourth reason that Greenberg's study
is limited (and this is the most restricting one) is that his method continues to
be what he praises as "analytic" (1974:25); it is fully logico-mathematical. In
his paper (1954) on synthesis, he is primarily quantitative; in his paper (1963) on word order, he not only counts, but explicitly employs set theory and symbolic logic (1974:17). As a result, while remaining close to linguistic data, Greenberg "... ends up with a series of specific universals that do not fit together as a coherent conceptual whole" (Comrie 1981:93).

2.2. Linguistics as Classification: Typology, Genealogy, Theory

2.2.1. The Divorce of Talk from People

In some respects, to "study Language" is a very curious thing to do. It is one of those enterprises that is impossible without the leisure afforded by knowing that one's belly is full now, and will be full tomorrow. The non-linguist, from an executive at Shell Oil to an Indian in the jungles of South America, has no need to "study Language" in order to use his own language. A common interpretation of the phrase "study Language" is that it means "to learn or to know many languages". The executive's first question to the linguist is "What languages do you know?"; and the field linguist either initially lies to his "language consultant" by saying that he wants to learn the informant's language, or else leaves the informant totally in the dark as to the purpose of his peculiar questions, in which case the hourly wage and mutual curiosity are what provide the tie between the two people. Yet both the executive and the Indian use their respective languages constantly; it is only the linguist whose belly is made full by making more words and perhaps pictures about Language. One can suggest other enterprises which might be thought to assume full bellies: art, perhaps, or music, or sports and games, or the contemplation of the nature of the cosmos. But there are striking
differences between these enterprises and the study of Language. Societies of all kinds have "produced" art, music, games and cosmic or religious contemplation; but it is only "civilized" societies that have "produced" linguists, or individuals who exclusively study "Language". In a non-"civilized" society, everyday life, or keeping one's belly full, does not permit artistic, musical, gaming or religious activities to be the daily pursuit of every person. There are temporary or permanent "specialists" in these activities, who play their roles on "special" occasions (appropriately enough). But everybody talks. There are none whose specialty is Language, except perhaps the sorcerer ... but even the sorcerer's main specialty is power, not Language.

The attempt to wrest words both from their everyday use for getting things done, and from the realm of the ancient magic, is one of many endeavours of the West's "advanced" culture. What recent Western philology and linguistics have accomplished with regards to this attempt has been touched on in my discussion so far. Enabled by the technology of writing, empowered by the methods of science and logic, and legitimized by the cultural dominance of a de-personalized, scientific, technocratic culture, the study of Language has come to divorce talk from everyday life and from people. A metaphysical Language is envisioned, and the odd discomfort that some specific language must be used to talk about Language is hedged. Thus Language is made reflexive. It is not certain that much either of use or of potentiality has been learned in the practice of making language reflexive in this way. Many marks on paper have been made, but whether any good has been furthered (other than the full bellies of the markers on paper) is not clear.
Among the linguistic approaches which have emerged from the reflexive study of Language in the last two hundred years, I have mentioned genealogical classification, typological classification, theoretical linguistics, and the Indo-European Grammatical tradition. I now wish to extend the proposal that there has developed not two, but three systems of linguistic classification (a suggestion of Philip Davis, personal communication).

2.2.2. Linguistic Theory and Literacy: Construction of a Classificatory Discourse

Theoretical linguistics may be conceived of as a classificatory attempt to tame Language, analogous to the classificatory attempt of typology and genealogy. Now there must be determined a basic unit of classification before one may classify. The method of genealogical study begins with consideration of arbitrary pairings of phonetic or phonological features with semantic ones, but ultimately the basic unit which it seeks to classify is the linguistic "family". Typological classification seeks a "type" (which has proved elusive) which it may employ as its basic unit. It is "grammar" that provides linguistic theory with its basic unit of classification.

Linguistics is impossible without literacy. The very conception of the linguistic enterprise as it is largely understood today may well have come with the devising of writing systems. Luria has indicated the connection between schooling (including the acquisition of literacy) and the ways of thinking and speaking which involve "the capacity to perform 'theoretical' operations of formal discursive and logical thinking" (1976:133). (See also Cole and Scribner 1974 and Tyler 1987 in this regard.) It is to be accepted that we live in a
literate culture; yet it is of value to attempt to become aware of the framework that literacy has given us to work within. Apart from the under-the-magnifying-glass interest in phonology, linguists have had to make tractable the flow of sound or (since the beginning of their discipline) its representation on paper. This they have called "Syntax". Syntax is divided into pieces in the course of the search for bits and pieces which Indo-European grammatical expectations have assured the linguist he is likely to find within a given language. Thus, there are stepping-stones in the formation of linguistic theory: armed with (perhaps unavoidable) expectations from Grammar, the linguist studies Syntax by means of writing down sound with the alphabet ("grammar"), and proposes therewith a Theory, or series of descriptive or explanatory statements; these then constitute Language.

Modern English grammar derives from Greek *gramma*, whose first meaning is 'that which is drawn or written, a written character or letter'; the plural form in Greek means 'letters, the alphabet'. Thus the suggestion that the reflexive study of Language is enabled only by literacy is corroborated by this etymology. It is of further interest to note that Greek *gramma* itself derives from the Greek verb *grapho*, whose first meanings are 'to grave, to scratch; to draw lines with a pencil, sketch, draw, paint; to write, inscribe'; but *grapho* also carries a legalistic overtone: 'to register, enroll; to write down a law to be proposed'. It is well-known that Grammar has largely served a prescriptive function for centuries; modern linguists attempt to extricate themselves from this function by democratically claiming to be descriptive. Though Modern English *grammar* ultimately derives from Greek, it comes into Middle English as *gramer* and into Old French as *gramaire*. In the Middle Ages, the word takes
on a new cast, in addition to the Greek one associated with writing: one associated with magic. Middle English *gramari(e)* has as its first meaning 'grammar, one of the liberal arts; also learning'; the second meaning is 'magic, enchantment'. In Old French the first meaning of *gramaire* is 'savant, astrologue, magicien'. Perhaps it is this magical aspect of grammar and writing which is currently undergoing a revival of sorts in some circles. For example, Bloch (1983:1-29), who entwines Levi-Strauss and Derrida, with Merlin as mediator.

The linguist’s Syntax evolves then from grammatical expectations. Modern English *syntax* comes through French and Latin from Greek *suntaxis* the primary meaning of which is militaristic: 'a putting together in order: of soldiers, a drawing up in order, a complete array: hence, order, arrangement: organization, system; a body of troops'; secondarily, Greek *suntaxis* has the meaning 'in grammar, the combination of words and sentences'; in addition, it (like Greek *grapho*) has legalistic overtones as 'a covenant, contract; a contribution, quota; a settled rate of renumeration'. The Greek noun itself comes from the Greek verb *suntasso*, meaning 'to put together in order, to draw up in order of battle, put in array; to arrange, organize, regulate, ordain, command'.

These etymologies intertwine well with a few suggestions I have made in this paper. Firstly, that the discreteness of *gramma* has enabled linguistic speculation. Secondly, that the organization, or the array, or the patterns of languages are ordained into "data" by a researcher, by a human "subject", someone who has been nurtured by educational and cultural expectations of "how things are". There is no objectivity in linguistic theory. Thirdly, that
"grammar", Grammar, and Syntax are actions; and they are actions of power and of command and regulation. Anthropologists today are realizing that ethnographies are not objective descriptions of the "Other" of foreign cultures, but rather are stories, dialogues, or instruments of domination. Perhaps it is time that linguists too became more overtly aware of the ultimately interpretive nature of their enterprise.

I have discussed the stepping-stones towards linguistic theory, "grammar" and Syntax; "grammar" is the unit which Theory classifies in some way; Syntax is a mediating activity. "Theory" has an interesting etymology too, and in combination with that of "grammar" and Syntax, things may become clearer to our eyes.

English theory comes through French and Latin from Greek theoria, which means firstly, 'a looking at, viewing, beholding, observing; especially the being a spectator at the public games'; its secondary meaning is 'of the mind, contemplation, reflection'; and thirdly, 'the sending of theoroi or state-ambassadors to the oracle or games'. Again note that the Greek word carries much more the notion of an activity than of a thing, and that the activity is primarily one of spectating (significantly, even mental activity has to do with seeing). Indeed, the Greek theoria is related to Greek theatron 'a place for seeing, a theatre', which comes from Greek thea 'a sight, a spectacle'. Thus Theory is a visual activity, and there is some performance to be watched. Linguistic Theory, one might say, is the spectating of "grammar", of the letters of the alphabet, the marks on paper, mediated by Syntax, the putting in order of "grammar". The method of linguistic theory as classification is literacy.
2.2.3. Typological and Genealogical Classification and the Scientific Method

Typological classification frequently still attempts to employ a method similar to the scientific one which has proved so successful for genealogical classification. However, language typologists have not been consistent or unanimous in their choice of what should form their initial starting point. The nineteenth century scholars typologized on the basis of the word; Trubetzkoy and Hockett did so on the basis of phonology; Sapir did so on the basis of a combination of morphological form and lexical and functional semantics; Martinet did so on the basis of prosodic features, and made proposals for the other "linguistic levels"; Greenberg did so on the basis of the order of "meaningful elements". As a result of this, there is a seeming uncertainty as to what is relevant for typology, and the popularity of typological classification among linguists has not been constant. Its "results" have been diffuse and as well have often strayed into "psychologisms". Genealogical linguists have not been so often led astray precisely because their method has defined which elements of language are relevant for analysis; the traditional scientific method employed requires discrete, tangible entities for manipulation, and such entities are readily available within the constructs of phonology and morphology. The excitement created by Darwin and the ideological climate which he helped to engender doubtless contributed to the legitimization of the genealogical metaphor, and to the spawning of "family" as a unit of linguistic classification. Though Trubetzkoy's work on phonological classification was pioneering and very influential for the development of phonological theory, the purpose of typological work is not to seek the deeper "form-feelings" of
languages of which Mathesius, Sapir and Boas spoke. The limited success of morphological approaches to typology has been already discussed. There is no chance, in genealogical linguistics, for any missing of the mark by the insidious creeping in of considerations of a nature which might be termed cultural, social, psychological, communicative, rhetorical or pragmatic.

Most theoretical linguists, too, have sought to keep their investigations clear and pure of these kinds of entangling considerations, as they objectify "Language". For example, Martinet (1962:138):

Linguists ... should not do what they are not trained to do and what might lead them into the realm of unverifiable hypotheses, ... and venture into the field of cultural history.

This statement is an example of the well-known restrictions placed on the study of language when it has as its goal the search for scientific truth. A scientific hypothesis is true only as long as it is not empirically falsified. Linguists who have attempted a linguistic typology have also been under this scientific constraint and, as I have said, have not often come up with clearcut "results", or universally testable scientific hypotheses. Alternatively, an attempt such as Greenberg's quantification of Sapir's morphological forms (Greenberg 1954) does have clearcut, replicable results, but their value and meaning is uncertain.

2.2.4. Typological Discourse: Marginality, Potentiality, and Time

In contrast to both theoretical and genealogical linguistics, typological study within the discipline has often been considered marginal. In part, this marginality may be a consequence of the isolation of the results of typology from the work of linguistic theorists and historians. The theorist Hjelmslev
(1970: 94-95), for instance, states categorically that

geometric and typological relationship are two quite different things that have nothing to do with each other .... The two divisions of the universe of languages have nothing in common. They are established on completely different bases and from completely different points of view.

He extols the investigation of typological relationship as "a rich and extremely important field, although still almost entirely uncultivated", but he remains adamant that "the resulting [typological] classification will cut across the genetic classification and be quite unrelated to it". The genealogical linguist Egerod (1980) invokes Hjelmslev's strict separation, and then investigates some current genetic dilemmas having great time depth, comparing genetic results with typological expectations stemming from Greenberg (1963b).

Repeatedly he corroborates Hjelmslev's separation; for example:

The Japanese language has a double origin .... The overall linguistic structure points to a strong influence from Altaic, but some typological criteria point towards Austronesian. We have a dilemma as far as genetic relationship is concerned -- and typological considerations cannot solve the dilemma (Egerod 1980:131);

and

Cham is genetically Austronesian ... typologically Austroasiatic. The two independent approaches give differing results, and neither can be used to support the other (Egerod 1980:133).

Egerod does ask "Is there something wrong with our concepts of genetic and typological relationship?" (1980:131), but he does not take the question seriously, and his tone is facetious. He merely affirms that postulations of genetic relationship should be based on sound correspondences, not on syntactic structure ("The mistake of assigning genetic affinity on the basis of sentence structure is less apt to occur nowadays" [1980:135]), and says "Let us in any case not throw out the methodology with the dilemma" (1980:131).
That is, he invokes the scholarly fraternity and methodology of traditional historical-comparative linguistics.

Linguistic historians such as Lehmann (1978c, d), Vennemann (1974), and Hawkins (1983) have sought to reduce the marginality of typology within linguistics by incorporating expectations generated by Greenberg (1963b) into genealogical linguistics, and it may be partially in response to them that Egerod is so insistent on his own position. Often, however, typology has been considered to be an unexplored territory of great potential for understanding the "real nature" of human language. Hjelmslev (1970:96), for instance, says that

An exhaustive linguistic typology is, in fact, the biggest and most important task facing linguistics. Unlike linguistic genetics, it has no regional limitations. Its ultimate aim must be to show which linguistic structures are possible, in general, and why it is just those structures, and not others, that are possible. And here it will come closer than any other kind of linguistics to what might be called the problem of the essence of language. Finally, it will prove to be superordinated to genetic linguistics, since only through linguistic typology can we hope to understand what laws govern linguistic change and what possibilities of change a given linguistic type implies.

Others have also as wistfully indicated the potential of the typological approach to the languages of the world (cf. Sapir's "underlying ground-plan").

An often mentioned difference between the genealogical and typological classification of languages is that the former involves history and the latter avoids the historical dimension. Greenberg (1974:10) gives as a primary characterization of typology that it is synchronic and ahistorical, opposing it on this basis to genealogical classification. It may be suggested that this characterization is somewhat simplistic, for a number of reasons.
It may be suggested that Greenberg's choice of the parameter of time as the basis of a comparison of the two approaches may be a choice that is an expected one. In the short run, he is simply echoing Saussure's dichotomy, no doubt; but in the long run, both he and Saussure voice expectations of their culture, and, more distantly, of their languages. To venture into "culture history" and "psychologizing" for a moment, it has been noted by some that Western cultures and Indo-European languages are relatively obsessed with the "concept of time" (e.g., Krishnamurti and Bohm 1985; Hall 1984). The exciting archeological discoveries of the nineteenth century contributed to the Euro-American cultural awareness of time and space which had already been heightened by the exploration and colonialism of the preceding centuries. Saussure's famous dichotomy of "synchronous" and "diachronic" was but a convenient labelling symptomatic of this increased time (and space) awareness and interest. Part of the thrill of historical linguistic work and genealogical classification stemmed from this increased awareness. It is a hallmark of many Indo-European languages that they are concerned with the grammatically precise locating of events within a temporal dimension, a concern that is certainly not shared with all languages. The extreme historicism of nineteenth century Western Europe was consonant with a cognitive expectation of "how things are", enabling such statements (even well into the twentieth century) as the following:

it is certainly less trouble to let the aura of past action pervade all that follows and determine the choice of past tense than to take advantage of the precise indication of the past contained in yesterday and dispense with tense endings. (Martinet 1962:140)

Greenberg's characterization of the genealogical and typological difference in this way is, then, but an example of the tendency in the Western imagination
towards an overriding concern with temporality as a mode of understanding. And Greenberg's characterization is typical of him in that it is more quantitative than qualitative.

The differences between typological and genealogical classification may stem from factors other than simply the presence or absence of temporal considerations within the methodology. I have discussed the results of linguistic investigation when guided by the use of scientific methodology, and by logico-mathematical theoretical expectations; I pointed out that both this method and these kinds of theories have restricted and framed both the questions asked and the answers to these questions which have been worked out.

In addition, the institutionalized research climate has at times not been as hospitable to typological research as it has been to genealogical, or to theoretical, research. To employ Kuhn's (1962) term, the latter have both provided "dominant paradigms": for example, historical-comparative studies in Germany in the 1870's, or the Bloomfieldian movement in the United States in the 1940's. Both genealogical and theoretical linguistics have, as Hymes (1974b) points out in his perceptive discussion of Kuhn's conception, constituted approaches whose "participants, and others around them, had a consciousness [sic] of a revolutionary change, and that there was indeed a paradigmatic community [sic]". However, Hymes notes also that "The paradigmatic community ... has never come to be equivalent to the whole of the discipline .... Each 'paradigm' simply has not succeeded in establishing complete authority" (1974b:10). Those interested in the typological study of languages have never constituted a dominant "paradigmatic community",
though there has been activity in this area of linguistic research for at least the last two hundred years. That Greenberg published his cross-linguistic research in 1963, in the midst of the dominance of the Chomskian "English-only" approach, is an interesting feature of his work, in this respect, for he was giving voice to a traditional, never absolute, strand of linguistic study. Thus, rather than being understood as different from genealogical classification because it is "minus time", while the latter is "plus time", linguistic typology as a research activity, or as a "discourse", might be characterized by its very marginality. Since it has not produced a dominant paradigm, its playing out has been more susceptible to the originality of individuals; also the nature of typological discourse has not been as polemic as that of either genealogical or theoretical linguistics, since it at no time has attempted to establish or to maintain authority.

2.2.5. Platonic Thought: Guiding Metaphor in Linguistics

Trnka (1929:70), a member of the Prague Circle (who in general reacted against the neogrammarian successes of 1870-1880) in his discussion of the "analytic method of comparison" (from which stems typological classification) and "comparative-historical grammar" (which stemmed from the comparative-historical method) suggests another basis upon which the two may be compared. He, too, suggests that the principal difference between the two methods involved "... is not, as Saussure believed, time, which is eliminated in synchronic study in contrast to diachronic study; the decisive point is the purpose of the study". Trnka (1929:70) characterizes synchronic study as that in which

one compares linguistic systems, whether or not they
represent successive stages of the same language or stages of related languages. In doing this, one proceeds in the same way as one does when comparing two or more paintings: one takes note of the colors, the drawings, the relationship of the parts to the whole; one discovers resemblances and differences; and one begins to see in recurring details the manifestation of determinative tendencies.

He contrasts this with the differing purpose involved in the historical method:

it is not a matter of comparing systems, or the particularities inside these systems, but rather of reconstructing the primitive image from which the examined copies proceed, or to place functionally equivalent traits of the original and of the copies into parallel.

I do not think that Trnka would disagree that both genealogical and typological classification initially involve comparison (cf. Ramat 1987:195). But Trnka makes a very interesting point about the differing purposes of these two methodological approaches. It has been remarked, in other discussions within linguistics, how the use of differing methods applied to the same phenomena will produce conflicting results (e.g., Hockett 1954; Egerod 1980). I have discussed varying attempts by typologists to apply (to the same phenomena) the scientific methodology of historical linguistics, and presented some of the varying results of such analyses.

But it is a different question that Trnka indicates. He is, as he says, examining the differences between the historical-comparative approach and the synchronic descriptive approach peculiar to the Prague school. Thus, he was not speaking specifically of typological linguistics; however, Mathesius' paper on "linguistic characterology" had been published the previous year (1928), and thus typological concerns had doubtless been discussed by the Prague group. It is not, then, out of place to extend his general remarks on the
synchronic study of language to the present discussion, particularly within the context of Greenberg's characterization of typological linguistics as ahistorical and synchronic.

To repeat, then, Trnka suggests that the purpose of the historical-comparative method is a matter of "reconstructing the primitive image from which the examined copies proceed", or to compare "functionally equivalent traits" of "the original and of the copies". In fact, it is only after this application of method (which results in the positing of etymons) that genealogical classification may occur. The practice of typological classification is similar to this in that the notion of "type" must be clear before the classifying may take place. This is the point at which the various proposed typological classifications have differed, as has been indicated. There had not only been indecision as to which "level" of language might provide a type, but confusion as to how to define "type" at all (for discussion, see Ramat 1987:19-29).

Keeping in mind Trnka's characterization of the goal of the historical-comparative method as being the search for the primitive source "image" of the 'copies" constituted by the data of existing languages, if one considers the meaning of "type", it can be suggested that typologizing may have a similar goal. Genealogical linguistics has among its successes provided us with etymologies, the contemplation of which can provide fruitful insights. Consonant with Trnka's characterization is the fact that English etymon comes (through Latin) from the Greek etumon, 'the true meaning of a word', neuter of etumos, 'true', 'real'; etumos is allied to Anglo-Saxon sod, 'true'. Both etumos and sod derive ultimately from the Indo-European participial form *essont
(from *es-, 'to be'), and thus carry with them the meaning "that which is".

English type is etymologically almost the inverse of English etymon. It comes (through French and Latin) from the Greek tupos, the first meaning of which is 'a blow'. It is most interesting to consider the secondary meanings of the Greek tupos in the present discussion. They may be grouped into three different semantic thrusts: 1) a copy or trace of something: 'the mark of a blow, the impress (of a seal), the stamp (of a coin), a print or mark of any kind'; 'figures or impressions wrought in metal or iron', 'image or statue of a man'; and, finally (interestingly, an auditory rather than visual trace) 'the effect produced on the ear by a blow, as the beat of horses' feet'; 2) tupos can indicate the potential of something: 'an outline, sketch, draught'; 3) it can indicate the origin of something: 'the original pattern, model, mould, type'.

The first and second meanings ('a blow', and 'the mark of a blow') of tupos indicate some kind of violent action; that English type is allied to Sanskrit tup, tump, 'to hurt' bears out this interpretation as well. Recall that in Tyler's study of classification systems he notes that "Typologies are not frequently used in folk classification, and are rare outside analytic contexts" (1978:275). Thus, the typological organization of knowledge is one which is most often found within the logical-scientific cultural setting. The variety and complexity of the world's languages, when subjected to typological analysis, have proved highly difficult to deal with. Attempted classifications have been so simplistic as to be meaningless, or so sweeping as to involve various kinds of forcedness or hedging in order to get the rambunctious data of the languages of the world to fit into pre-conceived analytic schemes. Sapir's suggestions alone do not involve this simplicity, or this forcedness: but neither is his purpose to
construct a classification.

The most expanded meaning of the Greek *tupos* is the idea of a copy or trace or mark of something else. Typological linguistics, as I suggested earlier, differs from theoretical linguistics in that it generally tries to stay "closer to the data" of language phenomena, rather than attempting to model those same phenomena. Genealogical linguistics, approaching the same phenomena of language, seeks, in Trnka's parlance, to postulate the "original image" (the etymon, the true, the real) of the "copies" which language data offer. Typological linguists have searched for a 'type' within language, upon the basis of which they may construct a classification. Etymologically, this would indicate that they are searching for *what* made the marks or traces that language data consist of; the insistence on this ever-elusive 'type' fundamentally precludes the possibility that language phenomena might be "all that there is", and that the 'traces' or 'marks' might be "true".

This same notion is implied for genealogical linguistics by Trnka's characterization of the comparative-historical method and by the etymology of English *etymon*. Thus historical linguists, too, are not content with language phenomena, but must look for something else, something more "essential" and "true", the "original image" of the "copies" they have in the data of languages.

This same notion is made explicit in theoretical linguistics in the concepts of *langue* and *parole* (Saussure), and "competence" and "performance" (Chomsky). It is the basis of phonology as well. One might say that it has been a guiding metaphor in linguistics. And Plato thought it all out long long ago!
An alternative goal to the one Trnka ascribes to genealogical linguistics is that it seeks ultimately to construct a classification. This, too, would seem to be the goal of typological linguistics. I have proposed that classification is the activity of linguistic theory as well. All seek to organize the plethora of the world's languages, to tame the perceived chaos of the multitude of tongues, and to do so by imposing neat classifications on them. As I have said, this impulse is highly different from that of Sapir, who was wary of classifications, and had the practical goal in mind of having a handy way of communicating about different kinds of languages. That later typologists have complained that Sapir was too dependent on form while at the same time saying that he was too psychologistic suggests that Sapir was not as taken in by his Platonic linguistic heritage as others have been.

2.3. The Dissolving of the Classificatory Approaches to Language(s)

This discussion has been concerned with the brief characterization of the three traditional subfields of linguistics: the genealogical, the historical, and the typological. These three "subfields" have generally been considered to be separate domains within the province of linguistics; the discourse of each has had borders which have traditionally been guarded. At times there have been border skirmishes and much linguistic academic discourse has been polemic in tone. This has been remarked on recently by Yngve (1986), who takes upon himself the task of providing an exhaustive set of rules of how the game of linguistic scientific discourse should be played, based on his training in the the discourse practices of science (his first academic field was physics). He is mistaken, I believe, in taking the cause of polemics in linguistics to be that linguists do not know how to "do science". Rather, it is partly that their
"object of study" (language) is not amenable to fruitful scientific analysis. It is also partly because the three subfields have developed in differing milieus.

I have provided a sketchy trace of the historical, methodological, and ideological backgrounds of these three traditionally separated subfields. Having this background before us helps us to recognize that what are apparently separate subfields, producing separate discourses, are actually separable precisely because of their historical development. In recent publications from all three of these classificatory "subfields" there have been proposals which indicate that dogmatic acceptance of these three classifications is being questioned.

2.3.1. Questioning by Linguistic Typologists

Payne (1987), for example, discusses (with reference to Papago narrative discourse) the problems associated with acceptance of the syntactic "typological tradition" of Greenberg and Hawkins (the tradition which came to take dominant syntactic orders of Subject, Verb, and Object as a "type" with which to classify languages). Payne mentions the problems associated with defining "the notions of subject and/or object" in some languages; with the "relevance of subject and object categories for constituent order" in other languages; and with "how the term 'basic' [with regards to word order] ... should be identified" (784-785). She concludes that "languages in which order is not based on syntactic role should simply not be forced into an order typology based on syntactic role" (783), and suggests that "order should be accounted for in some languages by syntactic factors, in others by pragmatic factors and in most languages by interacting factors of both types" (802). That is to say that the
notion of "type" as traditionally defined is irrelevant.

Ramat (1987:1-39) provides an excellent discussion of various recent developments in linguistic typological research, and reviews various definitions of "type". He himself proposes what he calls a "polythetic" typology rather than the "monothetic" traditional typologies. He writes:

Since 'type' is, moreover, an abstract model which is never realized completely, membership of a Type A or Type B language will not be decided in terms of a binary 'yes' or 'no'. In concrete linguistic reality, it will be necessary to speak not of 'mixed types', as has often been the case, but of typologically mixed languages (i.e., 'type inconsistent' languages). (1987:24)

The shift of interest from "language types" to "typologically mixed languages" is a significant one; this shift is also evident in Payne's (1987) position.

2.3.2. Questioning by Linguistic Genealogists

Ramat (1987) also relates the notion of "type" to historical linguistics, where it in part takes on the cast of what I have called "family" as a unit of classification in genealogical linguistics. He writes:

Here we shift into historical linguistics, because what comes to be considered a 'language type' is a set of languages which are either genetically related and share a series of features, such as Slavonic languages, or languages which became similar through contact, such as Balkan languages. If, however, we abandon the field of abstraction typical of theoretical models and move into complex historical reality, we can immediately see that the concept of type, which was reached with such difficulty, loses rigour and its boundaries are dangerously fudged .... (Ramat 1987:23)

There are others within genealogical linguistics who are de-emphasizing the notion of "family" and the reconstructions upon which "families" are postulated. For example, Austerlitz (upcoming 1989) voices skepticism
concerning the productivity and good of "long-range groupings" in genetic
classification (1). He writes:

It will not do simply to set up proto-forms. We must
remember that we are dealing with human beings who
face and overcome (or are overcome by) obstacles such as
mountains, the weather, and pestilences. (1989:9)

Austerlitz is hesitant to join those historical genealogists in linguistics who
posit long-range language relationships on purely formal manipulation of data,
without recognition of commonsense factors. Again, he writes:

I feel that the three branches just enumerated [Turkic,
Mongolic, and Tungusic] form a continuum (from West to
East, as described) which can be called a language family
but is not on the order of IE or Uralic. The answer to the
question [of long-range affinity] must be sought, I believe,
in social ... and economic ... organization. (1989:6).

Austerlitz here questions whether the understanding of "family" (as a unit of
classification), which has been gleaned from Indo-European and Uralic studies,
is indeed relevant for the languages he mentions. He also refers to Voegelin
and Voegelin (1985), who say that "Amerindianists had no choice but to work
with phyla ... [which were] often established by intuiting long-range [language]
groupings"; and that they have "remained saddled for generations" with these
classifications (Austerlitz 1989:8).

2.3.3. Questioning by Linguistic Theorists

Finally, there are linguistic theorists who question the notion of
"grammar". In this paper I have written of "Indo-European Grammar" and of
"grammar" as a unit of classification. Though I have separated these two for
the purposes of my discussion, the two really coalesce and intertwine more
than I have so far indicated. I have largely used the term "Indo-European
Grammar" to indicate the structure of the Indo-European languages; I have
used the term "grammar" to indicate the understanding of what a language is composed of, when conceived by Indo-European speakers within a principally analytic tradition. Thus "Indo-European Grammar" is encompassed within my use of the term "grammar". It is this traditional conception of "what language is composed of" that is being severely called into question by some theorists of language.

Hopper (1987) describes the traditional linguistic understanding of "grammar" in this way:

The assumption ... is that 'grammar' (in the sense of the rules, constraints, and categories of the language attributed to the speaker) must be an object apart from the speaker and separated from the uses which the speaker may make of it. That kind of grammar is conventionally understood to consist of sets of rules which operate on fixed categories like nouns and verbs, specify the forms of additive categories like those of case, tense, transitivity, etc., and restrict the possible orders in which words can occur in a sentence. Discourse, the actual use of language, is held to be in some sense an 'implementation' of these structures, or the way in which the abstract mental system possessed in its entirety by the speaker is realized in particular utterances. (1987:141)

Hopper notes that it is difficult to avoid (once trained!) the kind of thinking in which "the notion that structure pre-exists discourse", and suggests that in the attempt one may be caught in "the vicious circle of 'form-to-function-to-form'" (Hopper 1987:141). He proposes the term "Emergent Grammar" to encapsulate what is an essentially different approach to the study of language than that traditionally taught and written about:

The notion of Emergent Grammar is meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process. Grammar is hence not to be understood as a pre-requisite for discourse, a prior possession attributable in identical form to both speaker and hearer. Its forms are not fixed templates, but are negotiable in
face-to-face interaction in ways that reflect the individual speakers' past experience of these forms, and their assessment of the present context, including especially their interlocutors .... the term Emergent Grammar points to a grammar which is not abstractly formulated and abstractly represented, but always anchored in the specific form of an utterance. (Hopper 1987:142)

He notes that "The linguist's task is in fact to study the whole range of repetition in discourse, and in doing so to seek out those regularities which promise interest as incipient sub-systems" (1987:142); he notes that
real live discourse abounds in all sorts of repetitions which have nothing to do with grammar as this is usually understood: for instance, idioms, proverbs, cliches, formulas, specialist phrases, transitions, openings, closures, favored clause types, and so on. There is no consistent level at which these regularities are statable ...
(Hopper 1987:144)

Hopper says that he does not want to "abolish" the notion of grammar, but rather "to suspend it"; he proposes "the recontextualizing [of] the notion of grammar" (1987:147). In effect, Hopper is saying that "grammar" is not some thing which is attainable (much less innate!); nor is it an action which can be complete (and therefore represented). Rather, he writes:

Because grammar is always emergent but never present, it could be said that it never exists as such, but is always coming into being. There is, in other words, no "grammar" but only "grammaticization" -- movements towards structure which are often characterizable in typical ways. (Hopper 1987:148)

2.3.4. Dissolving Classifications

Thus, there is in some circles a movement away from the strictly classificatory approach to language which has been historically predominant within linguistics. This movement necessarily involves an abjuration both of the logico-mathematical approach and the scientific methodology which I have
described with regards to the classificatory approaches. This abjuration simply means that it is no longer important or interesting to seek to make abstract, discrete, finalizing classifications from within one's pre-determined subfield of diachronic, synchronic, or theoretical linguistics.
CHAPTER THREE

OBJECT-SUBJECT LANGUAGES: A STORY OF TYPOLOGICAL RARITY

3.1. Statistical Rarity of Object-Subject Languages

Whatever may be said at this time concerning Greenberg's (1963) introduction of linguistic typology based on word order, certainly twenty-five years ago it was a very appealing approach. Much work has been done by linguists along the lines suggested by Greenberg's paper; the dominant word order of a given language has become as much a "fact" to know about it as its genetic affiliation. Greenberg's raising of the Prague school concept of markedness from phonology to syntax, then, was a fruitful proposal.

If one glances down lists of languages of the world, noting the basic word orders, one can be struck by something which shall form the motivation for the language interpretations of the two chapters to follow. In by far most of the languages of the world, the basic word order is one in which the Subject precedes the Object: SVO, SOV, VSO. Pullum (1981) lists only thirty languages in whose basic word order the Object precedes the Subject: OVS, VOS, OSV. The excessive statistical rarity of Object-Subject (hereafter called "OS") languages is certainly striking, and the observation forms the basis of Greenberg's first "universal" (1963:77).

Chapter Four examines a language of one of these rare word orders: Hixkaryana (Carib), which has been classified as OVS in basic word order. Chapter Five is concerned with another of these OS word orders, VOS, and examines Tzotzil (Mayan) as an example of this word order. The approach to
language data employed in the interpretations of these two languages will now be introduced.

3.2. Semantics ("What Makes Sense") and Word Order Study

The language interpretations to be given in the next two chapters are based on the premise that a coherent understanding of word order is possible only if the problem is approached semantically. The term "semantics" will be used to indicate generally "the study of meaning". "Meaning", or "what makes sense", will be understood to emanate from a complex of interacting considerations. There are at least four considerations which are involved in this complex: lexical meaning, discourse structuring, discourse pragmatics, and cognitive or psychological considerations.

Brief remarks on each of these four considerations follow:

1) *lexical meaning* (language specific): Each recurrent "form" in a given language has a specifiable "meaning"; this "meaning" is rarely discrete and stationary, but may play differing cognitive (functional) or emotive or metaphoric roles at any given moment. (Historical linguistics has contributed to the understanding of this effect on a grander time-scale than that intended here.) That is, "prototypical" referential meaning is often of negligible importance when people talk. Various languages have various means for the indication of how any "form" is to be understood at any moment. For example, the Indo-European grammatical distinction between Nouns and Verbs is not one recognized by all languages; rather the notions of "thing-like" and "action-like" may be coded morphologically on various "forms" in order to assure their certain (though fleeting) understanding. Varying formal devices
are employed, then, to ensure that the instantaneous meaning intended is understood. There are gradations of various sorts available in "lexical meaning", and they are instantiated and made specific in any given utterance or discourse by their interaction with (at least) the other three considerations mentioned here.

2) discourse structuring (language specific): (A case in point of gradations of lexical meaning is, of course, the word "discourse", and the contemporary necessity of its specific definition whenever it is used, because of its current popularity. One must write "'Discourse' in the sense of ...".) There are a limited number of formal means available for languages to structure "discourses". By "discourse" here is meant any text or any spoken language situation, whether textualized or not.

Here the concern is not with "genres of discourse" (e.g., the scientific, the novelistic, the poetic, the taxonomic, and so on). Nor is the concentration on the relationship between speech and writing, or listening and reading. The cognitive, psychological, and cultural effects of literacy have been under exploration for some time; the previous discussion acknowledges that the possibility of linguistics relies on literacy, and probes some of the concurrent pitfalls. Writing has made still and linear the patter and music of speaking; speaking is the slowed down and discrete iceberg tip of the simultaneity and speed of thought, emotion, and intuition. The interest in these language interpretations are not why people talk, nor fully how they communicate. Rather, what is of interest here is the morpho-syntactic linguistic means with which human intelligence avails itself in languages in order to aid that fuller communication. The limited number of formal means (here called "discourse
structuring") which are available in languages are comprised of the following: prosody (intonation, and the use of silence); morphological elements (affixes, particles); and what has come to be called "word" order (by the latter is more properly meant that "order" of which Sapir writes). Thus, by "discourse structuring" is meant the "technical" means available to languages, and by means of which various semantics are signalled. Though the means available are limited cross-linguistically, the playing out of them in languages is idiosyncratic to each language (cf. Davis MS.b:57). Again, this consideration within any one language will be simultaneously intertwined with (at least) the other three considerations.

3) **discourse motivation** (individual and situation specific): Any time someone speaks, writes, or engages in a conversation, there is a purpose in doing so. Whence that purpose comes, or to what ends or how a language may be employed as a manipulative instrument of some kind (for good or for evil) is not the point of interest here. Rather the focus is on the linguistic means available to any speaker, listener, writer, or reader, means in which fluency is recognizable. In the fleeting context of a fluent discourse, "purpose" is tracked by the "technical" means of what here has been called "discourse structuring". In a fluent speaker of a language, these means are rather "unconscious"; fluency may be defined by the full internalization of these means. This is far from saying that there is an innate language device in the brain; rather, human intelligence is innate, and a variety of its workings may be seen to be at play in the morpho-syntactic repertoire of languages. These workings of intelligence are not confined to, or specific to human languages; on the contrary, they are workings which play out in other realms of human activity.
as well. The manner in which the means of discourse structuring are acquired by speakers of a certain language is that of habit. The workings are "internalized" and become "unconscious" by repeated use; it is a learning experience akin, to use a comparatively crude example, to learning to drive a car. A competent driver no longer "thinks about" how to drive a car, or how to negotiate traffic; a learner must concentrate on each of these technicalities. Repeated use and experimentation with a certain language confirms or thwarts certain expectations an interlocutor has about communicating with that particular linguistic configuration.

This is obviously somewhat different from the often consciously willed motivation involved in the questions of "why people talk" and "what can be effected by language"; it is very probable, however, that these two kinds of pragmatics differ in degree but not in kind. In the second-to-second playing out of any discourse, speakers constantly signal by linguistic means (among others, of course) what is to be understood as important. This sense of "importance" can be manifested in various distinguishable realms: in terms of information that is KNOWN and UNKNOWN to the interlocutors up to that point (GIVEN-NEW); in terms of what is being talked about (TOPIC); in terms of what is somehow new concerning either what is KNOWN and UNKNOWN, or what is being talked about (RHEME). (Following the practice of Davis [MSS.a, b, c], "semantic" terminology will be capitalized.) Again, this consideration of "discourse pragmatics" both interacts with and is simultaneously acting with the other three considerations mentioned here (as well as with the more conscious kind of pragmatics mentioned above, with proxemics, with "body language", etc.).
4) cognitive or psychological considerations (pan-human): Increasing attention is being paid by linguists to the findings of biology and psychology in the attempt to reach a coherent explanation for the diversity of linguistic form found in the languages of the world. That is, it is conceivable that such a coherent explanation for "how languages work" is not to be had by a lonely wrestling with linguistic data, nor by struggling to get linguistic data to fit into the mold of any specific, rigorous, and internally coherent linguistic theory which has yet been popular. Rather, Davis suggests (personal communication) that it is by means of looking indirectly at languages that their syntactic intricacies may well be able to be explained. There is no transcendental, objectified "Language", then, which may be gazed upon.

3.3. Approach and Clarification of Terminology

There are alternatives to this objectifying approach: for example, there are a multitude of situations in which language is used, and investigation of these situations can provide insights into "human" issues such as the expression of humor of various kinds, manipulation and survival, efficacy and performativity ("getting things done"), and play, and contemplation. There is an interplay between language, ideology, and the aesthetics of cultural expressions which is most interesting. However, the alternative approach employed in the following two chapters will be one which focuses on the the constitution of human psychological processes as principally described by psychologists and biologists, and as interpreted by Davis (MSS.. a, b, c), in his cross-linguistic studies of syntax.

Before embarking on the next chapter, some terminology must be
introduced in the sense employed by Davis. Of particular interest for what is to follow is Davis' use of "RHEME". This term was mentioned just above, in the discussion of "discourse motivation", and was described as follows: "'important' .... in terms of what is somehow new in the immediate discourse, concerning either what is KNOWN and UNKNOWN, or what is being talked about". The semantics of RHEME, then, functions independently both of the scale of GIVEN-NEW information, and of the semantics of TOPIC. Though it functions independently of these two semantic parameters, RHEME also may often overlap with them; permutations of "discourse structuring" in a given language permit one to distinguish both the independence and the overlappings of these parameters.

Davis describes the function of TOPIC as the organization of experience or knowledge that is in all cases KNOWN to the speaker/listener; RHEME he describes as functioning to organize experience or knowledge which is UNKNOWN to the listener (MS.b:49). While TOPIC functions to guide the listener where in his memory or knowledge to enter the content of a PROPOSITION (MS.c:37), RHEME functions to direct the listener's attention to some element of a PROPOSITION because it is "new" in some way.

It is necessary to distinguish the "newness" of RHEME from that of the "NEW" of the GIVEN-NEW scale. The latter scale concerns only the discourse status of information: whether or not a named PARTICULAR is IDENTIFIABLE or not, and if so, the extent to which it is. Thus "NEW" on this scale has a value which involves the introduction of an UNKNOWN and UNIDENTIFIABLE PARTICULAR. The "newness" of RHEME derives from something else, and though it may overlap with the introduction of NEW
information, it does not necessarily do so; Davis (MS.a:28) writes that

the content of 'identifying the portion that is to be focally attended' is very frequently motivated by some portion of the PROPOSITION being UNKNOWN or NONIDENTIFIABLE. But that is not the essence of RHEME; focal attention is.

RHEME may, for example, mark an element of a sentence as important because it either accords with or does not accord with expectations up to that point in the discourse; that is, it may mark change of some sort.

Davis calls RHEME a "selective device" (MS.a:28); he writes (MS.a:32):

The expression of RHEME in its frequent, varietal manifestation of 'what the listener doesn't know about what's being talked about' is forward-looking, the compelling drive of a conversation, its motivation.

While the "NEW" of the GIVEN-NEW scale is an "absolute" NEW, in that the PARTICULAR is UNKNOWN and UNIDENTIFIABLE, "the newness of RHEME is relational and relative to a network of knowledge that is the PROPOSITION itself" (MS.b:50). That is, RHEME marks something as "unknown-in-some-conversational-context" (MS.a:27); thus, for example, "a PARTICULAR is UNKNOWN with respect to its functioning within a PROPOSITION, e.g., as fulfilling some ROLE; and an EVENT is UNKNOWN with respect to PARTICULAR PARTICIPANTS" (MS.b:50).

This introduction to Davis' use of terminology is sufficient for now; more of his concept of language will become apparent in the language interpretations to follow, and will be explicitly discussed in Chapter Six.

3.4. Word Order Interpretation as a Narrative Exercise

The complex interaction of at least the four considerations just discussed is
such that a coherent understanding of how languages work is incompatible with linguistic modelling or formalization. This is partly because of the dangers of such simulacra in general (discussed in Chapter One), but also because these interactions are present both simultaneously and momentarily. That is, languages are more like actions than like things. However, even this analogy is insufficient, for languages are neither actions nor things. Perhaps a more illuminating analogy is one with particle physics: any utterance or discourse, like a particle in a chamber, pops in and out of existence, is intimately involved with simultaneity, and is "known" only by its "traces", or by what effects it has. For the linguist as for the physicist (and many others), it depends upon which question the researcher is seeking to answer at any moment as to which consideration will appear most relevant. Thus the (at least) four considerations which have been mentioned here continually overlap, and not one of them alone is capable of explaining "how languages work" (much less providing answers to some of the other questions to which allusion has been made).

The interpretations of word order which follow, therefore, are presented in a narrative form (which implies space, time, and motion), rather than in scientific form and diagrams (which imply only space). Again, it is fully acknowledged that the four considerations mentioned above (and probably more!) are active simultaneously; this is all the more reason to tell a story rather than draw a diagram or to make a model. Any story moves along, and takes time to listen to or to read, and takes place somewhere, whether orally in a physical setting, or whether in writing upon a page at a bus stop or in a library, but always in the minds and emotions of the teller and listener. It is in
the listening to many stories that an approximation to simultaneity is made. Formal models and diagrams fail in this because they are based on simultaneity: understanding in an instant, by means of spatial representation, rather than understanding through time, by means of a kind of repetition.

The first story to be told here, then, is that of Hixkaryana word order, and why it is that this order is such a rare one from the perspective of syntactic typology. In studying and writing about the Hixkaryana language, Desmond Derbyshire has brought to the attention of linguistic typologists a language whose syntactic word order (OVS) was thought possibly non-existent. He has studied the interface between the proposals of the Greenbergian typological "tradition" and Hixkaryana, and appears only partially satisfied. In his various publications on the language (1965, 1977b, 1979, 1985), Derbyshire has provided both a large amount of syntactic data and invaluable commentary concerning the discourse context of much of these data, which has greatly enabled the interpretation to follow.
CHAPTER FOUR

HIXKARYANA (CARIB): OVS

4.1. Hixkaryana and Word Order Typology

Hixkaryana is a member of the Carib family of languages, and in 1979 there were about 350 Hixkaryana speakers; they were living on the rivers Nhamunda and Mapuera in northern Brazil, about halfway between Guyana's southern border and the Amazon (Derbyshire 1985:xiii). Hixkaryana is the best documented of a number of purportedly OVS languages spoken in the geographic area which drains into the Amazon: "A circle drawn with Belem on its circumference and Manaus as its center would include the location of every object-initial language that we know of, living or extinct" (Derbyshire and Pullum 1981:211). Not all these languages are classified as Carib, and there are Carib languages which are not object-initial. (See Derbyshire and Pullum 1981 for detailed discussion.)

Derbyshire lived with the Hixkaryana "for various periods between 1959 and 1975, amounting to a total actual residence of about seven and a half years" (Derbyshire 1985:xvi). Derbyshire mounts substantial evidence that the "basic word order" of Hixkaryana is OVS. His work therefore necessitates modification of the statement of Greenberg (1963), that of the six logically possible orders of major constituents (SVO, SOV, VSO, VOS, OSV, OVS), "only three normally occur as dominant orders. The three which do not occur at all, or at least are excessively rare, are VOS, OSV, and OVS" (Greenberg 1963:76). This hypothesis of Greenberg led him to the "universal" claim: "In declarative sentences with nominal subject and object, the dominant order is almost
always one in which the subject precedes the object" (Greenberg 1963:77).

Derbyshire (1985:95) notes that typological studies since the publication of
Greenberg (1963) have made even stronger claims than Greenberg's; for
example, Derbyshire mentions Vennemann (1973), who, he says, misrepresents
Greenberg "by making his universal more absolute than it really was". Derbyshire cites Vennemann (1973:27): "'Greenberg observes that of the six
possible arrangements ... only three occur as the only dominant pattern of
declarative clauses, viz. those in which S precedes O .... This is readily
explained'.". After the publication of Keenan's (1976) work on Malagasy, which
showed that language to have a basic order of VOS, the claim that S always
preceded O was shattered, but this "did not seem to affect the general
assumption that object never occurs sentence initial" (Derbyshire 1985:96).

Derbyshire (1985:96) notes that Pullum made "the strongest statement of all
that an absolute universal of word order is valid", and quotes Pullum
(1977:269):

"Four basic orders, not three, are found: SVO, SOV, VSO,
and VOS. The other two logically possible orders, OSV
and OVS, do not occur at all, contra various allusions in
the literature on syntactic typology."

It is within the context of this Greenbergian typological discussion that
Derbyshire writes one of his articles (1977a) and one of his books (1985).
Firstly, Derbyshire seeks to establish that Hixkaryana is indeed OVS in its
"basic order"; secondly, he discusses the syntactic correlations and differences
between Hixkaryana and existing typologies (1985:105-144).

4.1.2. Evidence for OVS Word Order

Derbyshire's claim that Hixkaryana is OVS in word order has been
generally accepted; Hixkaryana is considered as the only documented, "attested" OVS language. Thus, and in the context of the succeeding interpretation of Hixkaryana word order, it is important to consider both the kinds of evidence Derbyshire presents for his claim, and how he defines the notion of "basic order".

First, then, three kinds of evidence are presented by Derbyshire (1977a; 1985:97-99) to substantiate his claim that Hixkaryana is OVS in syntactic order. The first kind of evidence is that of intuition:

Intuition alone may not carry too much weight ... It is not, however, entirely irrelevant, and my claim here is simply that through various periods of residence among the Hixkaryana totalling more than seven years, my intuitive feel has always been that the most natural order in sentences is OVS .... I have translated the New Testament with the help of many native speakers .... I have observed this usage in contexts as varied as normal conversation, personal narrative, folktale narrative, eliciting of data, sermons, and written letters, as well as the reactions of informants to my attempts to produce well-formed and natural constructions. While I have often heard sentences with a different order of constituents, OVS has always seemed to me to be the basic one.

Derbyshire gives the following series of sentences as exemplifying this basic order "in each of the major sentence types" (Derbyshire 1977a:593):

1) a. Transitive OVS
toto yonoye kamara
   man he-ate-him jaguar
   'The jaguar ate the man'

   b. Intransitive VS
   nomohtxowni horoto
   they-came spider-monkey
   'Spider monkeys came'

   c. Stative COMP-COP-S
   amotohari nehxakoni toto
   not-tiring he-was man
   'The man was not tiring'
d. Quotative OVS
roti itxempoko kekoni kurumyana
my-meat-food order-to-poison-it he-said buzzard-person
'Order (him) to poison fish for me, said the buzzard man'

e. Equative PS
oruen hokono omoro
clay-pot one-occupied-with you
'You (are) one who is making a clay pot'

The second kind of evidence presented by Derbyshire is that of a statistical nature, gleaned from "a random sampling of texts"; in this sampling, there were "twice as many clauses with postverbal subject ((O)VS) as those in which subject is sentence initial (S(O)V). Object occurred immediately before the verb in all except three cases of VO" (Derbyshire 1985:97). It is important to note that "it is common in Hixkaryana discourse to omit nominal subjects and objects (and other nominals) whose referents have been established earlier in the discourse" (Derbyshire 1977a:593), and that "such occurrences were included in the sampling" (Derbyshire 1985:97); this accounts for the bracketed "(O)" 's in the passage from Derbyshire just quoted. Derbyshire (1977a:594) lists the permitted syntactic orders as "VS, SV, OV, VO, OVS, SOV", noting that "All except VO are frequent". He summarizes "the statistical facts of importance" as follows (Derbyshire 1977a:594):

(i) the object [when not elided, EAC] nearly always occurs immediately preceding the verb, the only exceptions being the rare occasions when it immediately follows the verb;
and (ii) although the subject [when not elided, EAC] often occurs in sentence-initial position, it is more frequently in final position.

Concerning this statistical evidence, two points may be noted. First, as Derbyshire specifically says, "many clauses do not have either subject or object NP's" (1985:97); there are many sentences in which either the Subject or the
Object has been elided, "their referents having been established earlier in the discourse" (1977a:593); there are obligatory prefixes on the verb, which, in Derbyshire's interpretation, mark the person of both the Subject and the Object (1977a:593). Second, Derbyshire has chosen the Subject (rather than the Object, or any other element of the sentence) as the relevant parameter upon which to conduct his statistical evaluation, hence the bracketed "(O)"'s in the quotation to which attention has already been brought. Derbyshire is making a strong attempt to rectify the claim of earlier syntactic typologists who cast doubt on the existence of an OVS language. If one accepts that "O's, V's and S's", that is, syntactic elements, are the necessarily relevant parameters upon which to compare and to (perhaps) classify diverse languages, then Derbyshire's presentation succeeds: the "unmarked" order in the Hixkaryana sentence is OVS. However, it may well be that these syntactic elements are not those which will best permit an understanding of Hixkaryana word order, and that discourse-pragmatic-cognitive parameters will provide that understanding. This is in concurrence with, for example, Payne (1987:783), quoted above: "languages in which order is not based on syntactic role should simply not be forced into an order typology based on syntactic role".

Before further discussion of this, however, Derbyshire's third type of evidence, and his concept of "basic order" must be presented. The third type of evidence presented by Derbyshire (1977a:595-598; 1985:98-99) is what he calls "syntactic evidence", and it is of two sorts. First, he is interested in ruling out as "basic" "the only other common surface order, SOV"; he writes that two rules will account for all occurrences of this variation from OVS order (1985:98):
the obligatory question-word fronting rule...; and the optional emphasis fronting rule ..., which is applied under discourse-pragmatic conditions that can be stated with a fair degree of precision .... These rules apply to adjuncts (including indirect object) as well as subject, but with this significant constraint: not more than one constituent can be fronted in the same clause, so that if a subject is fronted there cannot also be a fronting of an adjunct constituent, and vice-versa. If SOV were assumed to be the basic order, there would be no explanation as to why an adjunct could not be placed before the S, in sentence-initial position, for questioning or emphasis.

Again, one can see that Derbyshire's main concern is to place Hixkaryana within the Greenbergian syntactic typology; the linguist's common desire to predict has been discussed earlier, and will not be further commented upon here, except to note that in the interpretation which follows, the redundancy of "obligatory rule" and the contradiction of "optional rule" dissipates, and "discourse-pragmatic conditions" provides a starting point rather than an afterthought. (Please note that this is not a criticism of Derbyshire, whose work is meticulous and has enabled the interpretation which follows; it is rather, a criticism of the theoretical tools he uses.)

The second piece of syntactic evidence presented by Derbyshire in his allegation of an OVS "basic order" concerns the status of discourse-initial sentences. These sentences, he writes (1985:98),

are generally (but wrongly, I believe) considered to be free of contextual conditioning and, for that reason, are often cited as instances where the basic order of constituents can reasonably be expected to occur. It is in this position where the Hixkaryana emphasis rule might be expected to apply most often, thus producing the marked order SOV, since highlighting of a newly introduced character often occurs in the case of the subject of the first sentence ... In fact, however, the unmarked OVS order occurs more often even in discourse-initial sentences .... the contexts make it fairly easy to see why some are highlighted and others are not ...
These discourse-initial sentences will be discussed in detail below, and in fact, will give strong evidence for the interpretation of word order which will there be given.

Finally, Derbyshire most recently (1985:96) defines the notion of "basic order" as "the order of constituents at the point where linear sequence is first defined in the grammar of a language"; earlier (1977a:592), he notes that "Surface orders ... frequently exhibit a number of variants, and it will be necessary to determine which of the different orders is the basic one", and concludes (1977a:592) the following:

> What is needed to establish the basic order, therefore, is an informed hypothesis about the grammar that will account for movements from that basic order, both structural (e.g. Wh questions) and stylistic movements (e.g. topicalization).

He subsequently (1985:97) adds the criterion of optimization of explanation:

> The ultimate test of any hypothesis about the basic order of constituents in a language is that it should permit the optimal generalizations in explaining the conditions under which other surface orders appear.

It may be noted that this is the same criterion is used by McCawley (1970) to argue that English is not SVO in word order, but rather VSO.

It will be suggested here that the grammatical constituents of Subject, Verb, and Object are not the parameters relevant for an understanding of Hixkaryana word order. Rather, semantic considerations of the kind mentioned in the previous chapter will prove more helpful. The suggestion being made here is that Hixkaryana is RHEME-initial; it is the semantic increment of RHEME which consistently co-occurs with the initial constituent in the Hixkaryana sentence, regardless of that constituent's syntactic or grammatical
status. All of Derbyshire's arguments for the predominance of a "basic" syntactic order of OVS can be seen as variations of this one statement. The reservations which were voiced concerning his statistical evaluation also dissipate, while corroborating his counts.

4.2.1. Hixkaryana: RHEME-initial

The basic grammatical order in a transitive sentence in Hixkaryana is OVS, with any other elements (including the Indirect Object) normally following the S:

2) kuraha yonyhoryeno biryekomo
   bow made boy
   'The boy made a bow'

3) kanawa yano toto
   canoe he-took-it person
   'The man took the canoe'

(1979:10)

Quotative constructions follow this same order obligatorily, with the direct quote initially, then the V, then the S:

4) kanawa yaryako rowti kano waraka
   canoe he-took-it my-brother he-said-it Waraka
   'My brother took the canoe', said Waraka

(1985:31)

Copular constructions too normally follow this order, as the Complement precedes the V, which precedes the S; any other element normally follows the S:

5) xamata ymo mkawo nehxakoni sarahora amnyehra
   rocky-island AUG on-top-of it-was manioc long-ago
   'The manioc was on top of the rocky island long ago'

(1979:35)

A copular sentence in a direct quotation exemplifies the normal OVS order twice:
6) itohra exko waraka yakoro kenano rohetxe rowya
not-going be Waraka with she-said-it my-wife to-me
"Don't go with Waraka", my wife said to me'

(1981:194)

The statement (by Derbyshire) that these copular and quotative constructions exemplify "OVS" word order must be based on the premise that there is something semantically similar about both Complements and direct speech, and that that "something", in turn, is somehow semantically similar to Direct Objects (hereafter referred to as "O") in sentence-initial position in Hixkaryana. (That direct speech and Direct Objects are semantically similar may have something to do with the fact that the only verb introducing direct speech is -ka-, which means 'to say', but also means 'to do': thus "to say something" is "to perform an action".)

4.2.2.1. Sentence-initial Position: NEW Information

The formation of questions and answers in a language represents an important diagnostic in the study of RHEME, since it demonstrates RHEME's most obvious and prototypical manifestation. In Hixkaryana, NEW information does not necessarily occur sentence-initially (bearing RHEME), but in questions and answers, as will be seen, the semantics of NEW information does overlap with the semantics of RHEME, as evidenced by the utilization made of sentence-initial position.

4.2.2.2. Wh- Questions and Their Answers

Derbyshire (1979:6) notes that information question-words obligatorily occur initially. Consonant with the notion that sentence-initial position is reserved for the expression of RHEME, answers to such information questions
have the requested information occurring initially. Consider the following pairs of questions and answers:

7a. onok yowtì yakoro mitano
    who brother-of with you-went
    'With whose brother did you go?'

    b. waraka yowtì yakoro itono
    Waraka brother-of with I-went
    'I went with Waraka's brother'

8a. eten ke haxa ryhe wewe yamano
    what with CONTR EMPH tree he-felled-it
    'With what exactly did he fell the tree?'

    b. yawaka ke haxa ryhe wewe yamano
    axe with CONTR EMPH tree he-felled-it
    'It was with the axe he felled the tree'

9a. onok min yaye momokno
    who house-of from you-come
    'Whose house have you come from?'

    b. rowti min yayo momokno
    my-brother house-of from I-came
    'I have come from my brother's house'

10a. henta minomno
     where you-left-it
     'Where did you leave it?'

    b. tuna yohoye inomno
     water above I-left-it
     'I left it at the river bank'

(1979:11)

It may appear from examples (7)-(10) that semantically sentence-initial position is reserved for NEW information, and with this series, that is indeed the case. But, as will be seen below, non-NEW information can also occur sentence-initially, and NEW information does not always occur in that position. However, there are two pertinent remarks that must still be made concerning the formation of questions and answers.
One is to note that only one sentence constituent can be questioned at a
time, so that the following sentences are unacceptable:

11) *onoki, eten xe naye
    who what desire he-is
    'Who wants what?' (1979:12)

12) *eteni, onok yonkukmeno
    what who it-baffles-him
    'What baffles who?' (1979:12)

The impossibility of having two elements questioned (and thus obligatorily
occurring initially) is analogous to the situation which Derbyshire (1979:71)
notes with regards to the "fronting" of constituents: only one constituent may
be "fronted" at a time. The limitation to a single constituent which Derbyshire
emphasizes in both information questions and "frontings" (both of which
concern sentence-initial position) give support to the present claim that it is
only this single position which is reserved for the expression of RHEME.

The second remark about questions and answers is the following: with
questions involving the O the questions follow the obligatory sentence-initial
positioning of the question-word; what is significant is that the answers to
these questions show no unusual word order, and in fact reflect what has been
described by Derbyshire as "basic" (that is, they are O-initial). Consider the
following, in which the O’s are questioned:

13)a. onoki ryhe menyo
    who EMPH you-saw-it
    'What did you see?'

    b. anaro horoto wenyono
    other spider-monkey I-saw-it
    'I saw another spider-monkey' (1985:62)

14)a. onok haxa ryhe menyo


what CONTR EMPH you-saw-it
'What was it you saw?'

b. maw yonyhera wahko; horoto haxa ryhe wenyO
howler-monkey not-seeing I-was; spider-monkey CONTR EMPH I-saw-it
'I didn't see howler monkeys; it was spider-monkeys I saw'

(1979:10)

In these information question-words and answers, then, there is again evidence of the semantics of RHEME being operable in sentence-initial position. The significance of the occurrence of O's initially (when it is they who have been questioned) follows the pattern of NEW information occurring initially (because they answer these questions); it also suggests that there is something semantically similar in Hixkaryana in the interpretation of NEW information (the obligatory initial placement of question-words, and their answers' initial occurrences), and with that of O's (because they occur initially both in response to these questions, and in "normal" word order). It is the semantics of RHEME which is being played out; both NEW information and O's are compatible with the semantics of REMOTEness that is associated with RHEME.

4.2.2.3. Copular Wh- Questions and Their Answers

The manner in which Hixkaryana expresses (English) copular constructions will further elucidate the function of sentence-initial position. As has been noted (cf. examples [5] and [6]), Derbyshire bolsters his claim that Hixkaryana is OVS by pointing out that the "basic" order in copular constructions is "Complement - Copula - Subject - Adjunct" (Derbyshire 1979:35); that is, he sees Complement-initial as somehow equivalent to O-initial. This observation will be shown to be consonant with the present
claim that Hixkaryana is RHEME-initial.

Questions and answers with (English) copular constructions reflect the same pattern in Hixkaryana as those which question O's (cf. examples [13] and [14]); question-words are obligatorily initial, and in answers the requested element comes first: both reflect "normal" OVS order.

"Identity" questions and answers in Hixkaryana are expressed with no V ("equative constructions"); note that NEW information occurs sentence-initially:

15) a. isoken komo mokyamo
    how-NOMLZN COLL those
    'What kind of people are they?'

    b. anhi komo mokyamo
    bad-one COLL those
    'They are bad people'

(1979:7)

16) a. onokyana omoro
    who-people you
    'Which tribe are you?'

    b. hexkaryana uro
    Hixkaryana I
    'I am Hixkaryana'

(1979:7)

17) a. eten oni
    what this-thing
    'What is this?'

    b. wayamakasi moro
    comb that-thing
    'That is a comb'

(1985:61)

18) a. onok yowti komo ryhe mokyamo
    who brother-of COLL EMPH those
    'Whose brothers are they?'

    b. waraka yowti komo ryhe moxamo
    Waraka brother-of COLL EMPH these
'These are Waraka's brothers'

"Non-identity" questions employ a form of the Hixkaryana verb -exe- 'be'; again note the "normal" sentence-initial placement of NEW information:

19)a. henta naxowi biryekomo komo
    where they-are child COLL
    'Where are the children?'

    b. romin yawo natxhe
    my-house in they-are
    'They are in my house'

20)a. onoki hyawo mehxako
    who with you-were
    'Who were you with?'

    b. waraka hyawo mehxako
    Waraka with I-was
    'I was with Waraka'

This section has examined the placement of NEW information when it occurs sentence-initially; the examples have involved the formation of information questions and answers. Two conclusions can be drawn: firstly, as is evidenced from variations in word order in response to specific questions, the semantics of RHEME in Hixkaryana may overlap with the introduction of NEW information; secondly, since O's and Complements when they have been questioned occur sentence-initially, the very relevance of the Greenbergian notion of a syntactic, "basic", "unmarked" order is called into question. As Derbyshire observes, whenever the O is nominally expressed (a prerequisite for determining such syntactic order) it is de facto in a position of emphasis.

4.2.3.1. Sentence-initial Position: Alternate Word Orders
Though the "basic order of [grammatical] constituents is OVS, with indirect object and adjunct normally following the subject" (Derbyshire 1979:40), other orders occur frequently as well. This section examines constituents that can occur sentence-initially in the language, and will provide further evidence that Hixkaryana is consistently RHEME-initial.

4.2.3.2. Subject-initial

The S may occur initially, resulting in an ordering of SOV; in such transitive sentences having no nominal elision, the English gloss is often a cleft-sentence:

21) *okomkurusu biryekomo heno yoskeko*
   
   *bushmaster child dead it-bit-him*
   
   'It was a bushmaster snake that bit the child'
   
   (1985:31)

22) *waraka yawaka yokheko rohyaka oroke*
   
   *Waraka axe he-sent-it to-me yesterday*
   
   '(It was) Waraka (who) sent the axe to me yesterday'
   
   (1977a:595)

The grammar of the English gloss sets the S in a position of emphasis within the English sentence; this emphasis in Hixkaryana is accomplished by having the S occur sentence-initially. Thus, when the S of a Hixkaryana sentence is to be interpreted as RHEME, it occurs sentence-initially rather than in final position. It is surmised that (21) and (22) are the only (non-elided) ways to answer the questions 'What bit the child?', and 'Who sent the axe to me?', respectively.

RHEME in English is marked by the varying placement of sentence accent, and by cleft-sentences (elements of "discourse structuring"). Thus, it may be construed that, in a pair of Hixkaryana sentences such as the following
(though they are glossed identically by Derbyshire), sentence accent (RHEME) in English has been shifted, in the manner marked:

23) a. kamara nahosiye
    jaguar it-grabbed-him
    'The j á gu ar grabbed him'
    (1985:32 [my accent])

b. nahosiye kamara
    it-grabbed-him jaguar
    'The j á gu ar grabbed him'
    (1985:32 [my accent])

Legitimation of this difference in the placement of English sentence accent (RHEME) may also be had from consideration of intransitive Hixkaryana sentences in which the S bears RHEME, and therefore occurs initially (or, as Derbyshire says, in which the S has been "fronted for emphasis"):

24) waraka haxa nehurkan ao asama yawo
    Waraka CONTR he-fell trail on
    'It was Waraka (not someone else) who fell on the trail'
    (1985:74)

25) kaywerye haxa nomokno
    Kaywerye CONTR he-came
    'It's Kaywerye (not the other) who has come'
    (1979:71)

Sentences (21), (22), (23a), (24) and (25) provide examples of what Derbyshire calls "emphasis" or "highlighting" being accrued the S; that is, they provide examples of variations in word order in response to the obligatory sentence-initial expression of RHEME in the language. Sentence (23b), then, provides an example of the verb bearing RHEME, if my placement of sentence accent in English is acceptable; the S is in "normal", final position and the O is elided.

4.2.3.3. Verb-initial
It has been mentioned that Derbyshire notes that VO orders do occur, although he says they are "rare". According to the interpretation being put forth here, these orders would provide examples in which the Verb bears RHEME, since it occurs sentence-initially. The Verb in these examples must also have some sort of semantic similarity to the O's (and S's and other elements) which more usually occur in this initial position.

Derbyshire (1977a:594) writes that "it is difficult to be sure" of conditioning factors, since VO order is so infrequent, but that it seems to hold true that it [VO order] is permitted only under the following conditions: (i) where the subject is first or second person and is marked only in the verb prefix; and (ii) elsewhere if the object is a morphologically and semantically complex item.

In a later publication (1985:76) he has not much altered this statement of occurrence; and he does not give any motivation for this "optional movement" of the O to the right of the Verb:

There are two conditions ... : (i) where it cooccurs with a I, I+II, or II subject; and (ii) where it is a "heavy" construction, and where the rightward movement also involves dislocation ...

In the few scattered examples of VO order given by Derbyshire, there is only one which is Verb-initial, and which does not involve any right-dislocation (the latter is characterized by a preceding pause, and a separate intonation pattern; it is indicated here by a comma):

26) texhtxe kana heno let’s-go-bring-it fish QUANT ‘Let’s go bring some fish’

(1977a:594)

Note both that the O is indefinite and plural, and that the EVENT is unrealized (future exhortation). There are two examples of VO order with
preceding Adverbials (of Time and Place, respectively):

27) awanaworo ipatxowi tehtxe kana heno
tomorrow let's-go let's-fetch-them fish QUANT
'Tomorrow let's go and fetch the fish'

(1979:78)

28) isna taryatxow hami katxhonano
to-there we(INCL)-take-it DEDUCT goods
'We must take the goods there, evidently'

(1985:76)

Note that in both of these examples, the O is plural, and that the EVENT is again unrealized: in (27) the future nature of the exorted EVENT is emphasized, as awanaworo, 'tomorrow' occurs initially, and bears RHEME; in (28), again the EVENT is future exhortatory, and the speaker has deduced (hami, 'DEDUCT') the place (isna, 'to-there'), the latter occurs initially, bearing RHEME.

All the other examples of VO order involve right-dislocation; in only one does the O come in the same phrase as the (initial) Verb; again the EVENT is unrealized (and is marked as "hearsay"), and the O is indefinite and plural.

29) itxemko ti kana, epepe yoti
poison-it HSY fish, my-older-brother meat-of
'Poison fish, food for my brother'

(1985:77)

The suggestion is made here that O's are "prototypically" definite or specific in reference (that is, SPECIFIC NEW); on the rare occasions when it is indefinite or non-specific (NONSPECIFIC NEW), the O occurs postverbally. That the unrealized nature of the EVENT may not be the relevant point here can be surmised from the following examples, in which the EVENTS are also future exhortatory; however, note that the O's are definite and specific in reference:
30) kanawa xenyko
canoe see-it
'Look at the canoe'

31) kyakwe wohra tesnye
toucan not-shooting let-us-two-be
'Let's not shoot the toucan'

Further evidence for this possibility comes from sentences such as (32) and
(33); the coordination involved in the O's of these examples could have been
expressed by a paratactic sequence, but instead the more definite/specific part
of the O is sentence-initial, and the indefinite/nonspecific part of the O is
postverbal:

32) kurahatho hnankaye waywa heno komo
bow he-laid-it-down arrow QUANT COLL
'He laid down the bow and set of arrows'

33) hakrya wotxowni ha koso heno komo
peccary they-shot-it INTENSFR, deer QUANT COLL
'They shot peccary and some deer'

The following sentences, which Derbyshire offers as VO examples, are not
really so in the Hixkaryana; it is only the requirements of English that
necessitate their gloss to be transitive, thus appearing as VO examples.

34) nahohsaxkon hati, amryehxahotho hati
he-used-to-grab-them HSY, one-who-had-gone-hunting HSY
'He used to grab them, anyone who had gone hunting'

35) wenyhoryetxehkan ha iro ha
I-finished-making-it INTENSFR, that-thing INTENSFR
'I have finished making that thing'

Note, however, that in (34) and (35), the part of the sentences before the pause
are complete in themselves, and could stand alone; the right-dislocated phrase
in each of these sentences further assures the identifiability of the elided (because GIVEN) O. Thus they are not examples of VO order, but rather can serve to elucidate the function of right-dislocation in the language; this will be discussed more in the next section.

It was suggested earlier that in verb-initial sentences, RHEME must co-occur with the verb. However, in the (rare) VO orderings just discussed, what seems to be determinant is the nature of the O itself, rather than real "emphasis" being accrued the verb. There is another manner of emphasizing the verb, which will be mentioned here for the sake of completeness. There is a class of words in the language which Derbyshire (1979:79-80; 1985:23) refers to as "ideophones", and which he describes as follows (1979:82):

The ideophone is a noninflected, onomatopoeic [?, EAC] word that denotes an action that is normally expressed by a finite verb form. It is the only class of word that cannot be followed by a postpositional particle. It functions normally as a distinct sentence constituent, carrying the same meaning as that contained in the finite verb of that sentence (i.e. the basic part of the verb meaning, not including person, tense, etc., which are not part of the meaning of the ideophone) .... The ideophone may be a single morpheme (e.g. kritik 'action of making with the hands') or a sequence of reduplicated forms (e.g. sih sih sih sih 'action of walking'); .... Ideophones also show some variations from the normal phonology ...

He notes that these ideophones "occur very frequently in most types of discourse" (1979:79), and that they "Usually occur sentence-initial or sentence-final" (1979:80). The examples which he gives all are sentence-initial:

36) to , kay hati , kamara
dropping-down, he-did-it HSY , jaguar
'The jaguar dropped down (from a tree)' (1979:80)

37) bo , irakataworo
action-of-falling-into-hammock, in-the-middle-of-it
'He slept at night during the trip'

(1979:80)

38) txe, nahohsatxkon hati, amryehxahotho hati
grabbing-action, he-grabbed-them HSY, ones-that-went-hunting HSY
'Grabbing, he (the jaguar) used to grab them, the ones who had
gone hunting'

(1979:80)

39) ipo, nahatakaye owto hona
emerging-into-open, he-came-out village to
'Coming into the clearing, he arrived at the village'

(1985:23)

It seems clear that these ideophones emphasize the verb, in a way that is
consistent with the semantics of sentence-initial RHEME. It may be that
ideophones serve to make the action of the verb into a TOPIC (see "TOPIC as
REME", below), but further study of extended texts would be necessary to
ascertain this.

4.2.3.4. "Discontinuous Constructions"

Sentences (34) and (35), considered as examples of right-dislocation, begin
also to overlap with what Derbyshire calls "discontinuous" constructions
(1979:77). Since the O is of special interest in the present discussion, the
examples given here will mainly involve "discontinuous O's"; but such
discontinuous constructions occur regarding other grammatical elements too
(the S, Adjuncts, Indirect Objects, and Complements). Derbyshire (1979:78)
says that such "discontinuous" O's (and Complements) "serve the purpose of
either coordination or greater degree of specification" (see also Derbyshire
1985:77). It is important to note at this point two characteristics of
Hixkaryana to which Derbyshire draws attention: first (1975:45),

There are no formal means in the language for expressing
coordination at either the sentence or phrase level, i.e. no simple equivalents of 'and', 'but' and 'or';

second (1985:6), that

Nouns are not formally divided into classes or genders, and none of the following categories are morphologically marked: definiteness, indefiniteness, and genericness.

The following examples of a "discontinuous" O does seem to involve a "greater degree of specification", as Derbyshire says:

40) roti yimyako, kana
     my-meat he-gave-it, fish
     'He gave me meat, that is, fish'

(1985:77)

Sentence (40) is analogous to (41) and (42), in that the right-dislocated element provides further specification of the sentence-initial constituent:

41) kanawa weny o, asako
     canoe I-saw-it, two
     'I saw two canoes'

(1985:26)

42) owto hona kahatakeko, anaro owto
     village to I-came-out, another village
     'I arrived at another village'

(1985:26)

Sentences (41) and (42) may be contrasted with the following two, in which the sentence-initial constituent is apparently NEW information; (43) and (44) may be surmised to be possible answers to the questions 'What did you see?' and 'Where did you arrive?', respectively:

43) asak kanawa weny o
     two canoe I-saw-it
     'I saw two canoes'

(1985:26)

44) anaro owto hona kahatakeko
     another village to I-came-out
     'I arrived at another village'

(1985:26)
It must be construed that since there is nominal expression of the O’s in (41) and (42), that their referents are not of GIVEN discourse status (otherwise they would be elided). It is possible that they are of NEW status; but notice that there is also some NEW information in the final part of the sentence (in [41] and [42]): *asako*, ‘two’ and *anaro*, ‘another’; however, within their own "part", these elements are still initial. The NEW-ness that is in right-dislocated position serves to make more specific information that is KNOWN and just mentioned (that is, the O’s *kanawa* and *owto*, respectively). This, according to Davis, is the discourse function of TOPIC: to organize KNOWN information. That right-dislocated elements have to do with discourse TOPICs is further suggested by sentences such as the following:

45) tkatxefaranke xarha nehxtxkon hati,
    having-machetes also they-were HSY,
    iyeorye komo rma hati
    their-weapon COLL SAME-REF HSY
    'As to weapons, they had machetes also'

(1979:75)

46) isna rma txko tyufa nkekoni,
    to-there SAME-REF DIMIN spitting-action he-did-it,
    oseryehri
    his-being-afraid
    'As an expression of his being afraid, he was spitting into the little pot'

(1979:76)

There are a number of sentences involving right-dislocation and "discontinuous" grammatical elements, in which PARTICULARS are referred to rather obliquely (by pronominals or demonstratives) in sentence-initial position, and these PARTICULARS are then named in the right-dislocated phrase. For example:

47) oni enyhoryeko, epepe yahoni
this-thing make-it , my-older-brother seat-of
'Make this (stool), a seat for my older brother'

(1979:41)

48) ito tonyetxhe , owto ho
there we-eat-it , village at
'We will eat it there, at the village'

(1979:41)

49) oskeno tho yosahtotxowni hati saraho tho
one-thus DEVLD they-made-a-place-for-it HSY manioc-shoot DEVLD
'They made a place for that thing, the manioc-shoot'

(1979:78)

50) rokatxho wanimno , waywi komo, kuraha komo, mukawa komo
my-thing I-picked-it-up , arrow COLL, bow COLL, shotgun COLL
'I pick up my things, the arrows, bow and shotgun'

(1979:11)

It is probable that these are examples of what Derbyshire (1977b:178)
calls "one pattern of redundancy" in Hixkaryana, and which "functions in
relation to the relative strength of identification of a participant". Derbyshire
describes a "ranking concept" used by Grimes (1975:92-93) which

states that there exist series of identifications of the same
participant, not necessarily in contiguous clauses, in
which no identification is stronger than the one before it.
The scale of strength goes from proper names to explicit
descriptions to common nouns to nouns used generically to
pronouns, and from there to reference without
identification. (Derbyshire 1977b:178)

Derbyshire notes, however, in contrast to this "ranking concept", that in
Hixkaryana, identification of PARTICULARS most often works the other way
around; that is,

the ranking can also go the other way, beginning with the
weakest type of identification, the person-marking verb
prefix, and following with ellipsis, pronoun, general noun,
descriptive noun, kinship term, personal name, in that
order of increasingly strong forms of identification. This is
the most common form of referential redundancy in
Hixkaryana, and its effect is to narrow down to a desired
level of specificity, thus avoiding potential ambiguities. It
occurs both within a sentence and across sentence
boundaries. (Derbyshire 1977b:178)

This observation by Derbyshire can add to the understanding of Hixkaryana word order. The semantics of sentence-initial position are very powerful; there is a correlation between this position and the oblique introduction (MENTION) of NEW participants. That is, with regards to the introduction of participants, the language is rather suspenseful, for it identifies PARTICULARS only gradually, ASSERTing their identification near the end of a sentence or section of discourse. Derbyshire (1977b:176) suggests that "the major function of redundancy is to slow down the rate of introduction of new information", by means of its dispersal across phrases and sentences. This is the corollary to the suggestion being made here that Hixkaryana is RHEME-initial. If: a) what is to be interpreted by a listener as somehow new or unknown-in-the-conversation-at-hand (RHEME) consistently occurs sentence-initially; and b) this includes the identification of PARTICULARS; and c) there is no morphological mark of definiteness/indefiniteness; then, the redundant "sentence clusters" that Derbyshire describes (1977b) and his reference to the desire of "slowing down the rate of introduction of new information" can be explained.

4.2.3.5. Object-initial

It remains to see how Hixkaryana would express a sentence in which the O is emphasized, and bears RHEME. It has already been noted that both the S and O are elided when recoverable from the preceding context, though they are tracked to some extent by prefixes. Derbyshire is more specific than this, however, and notes (1985:76) that the O is elided except when "highlighted" in
some way:

I regard the normal, clause-initial position of direct object as the case of unmarked emphasis, since an object NP only occurs when there is some kind of highlighting or focusing of the constituent.

Thus, the O is obligatorily elided except when it bears RHEME; in this case it occurs nominally, and occurs in sentence-initial position; this, of course, results in the "normal" OVS word order. The question then of how Hixkaryana would express what in English is a "clefted-object" (e.g. 'It was the child the snake bit') would presumably be identical to the expression of what in English is an "unmarked" transitive order (e.g., 'The snake bit the child'): a "normal", OVS sentence.

Consider the following two sentences, in which constituents other than the S (the Indirect Object and Temporal, respectively) have been "fronted" (and cf. [22], in which the S of this sentence occurs initially):

51)a. rohyaka yawaka yokheko waraka oroke
    to-me axe he-sent-it Waraka yesterday
    '(It was) to me Waraka sent the axe yesterday' (1977a:595)

    b. oroke yawaka yokheko waraka rohyaka
    yesterday axe he-sent-it Waraka to-me
    '(It was) yesterday Waraka sent the axe to me' (1977a:595)

Note that [22], [51a], and [51b] have been glossed in English by cleft-sentences; only [51c] of this series does not necessitate a clefted gloss in English:

    c. ("basic order")
    yawaka yokheko waraka rohyaka oroke
    axe he-sent-it warka to-me yesterday
    'Waraka sent the axe to me yesterday' (1977a:595)

Comparison of these four sentences, along with Derbyshire's comment (quoted
above) that he regards "the normal, clause-initial position of direct object as the case for unmarked emphasis" give support for the probability that an appropriate alternate gloss for [51c] would be 'It was the axe Waraka sent to me yesterday'. Derbyshire's comment concerning this series of sentences involving "fronted" constituents is that this "fronting" "marks a variety of functions, including emphasis, focus, and cohesion" (Derbyshire 1977a:595). All of these "functions" are consonant with the present claim that Hixkaryana is consistently RHEME-initial. Derbyshire's "emphasis, focus, and cohesion" all are possible descriptions of the semantics of RHEME; each of them draws the listener's attention to something important and "new" in the immediate discourse context.

Though the English glosses in the following three sentences do not overtly reflect the RHEME accrued the initial constituent, it is to be surmised both that sentence accent in the English would fall on the "fronted" Hixkaryana constituent, and that an alternative English gloss would be comprised of a cleft-sentence:

52) watma ke netahetxoni toto heno homo club with they-killed-them person dead COLL
The people used to kill them with clubs' (1985:19 [my accent])
'It was with clubs that the people used to kill them', EAC

53) yayhi hoye netmatxowi waraka tapir from he-gave-them-meat Waraka
Waraka provided them with meat from the tapir' (1985:19 [my accent])
'It was from the tapir that Waraka provided them meat', EAC

54) rowti hokyamo ekeh me natxhe atunano wya my-brother children-of sick-one DENOMLZR they-are fever by
'My brother's children are sick with fever' (1985:32 [my accent])
'It is my brother's children who are sick with fever', EAC
Thus examples of non-OVS orders in Hixkaryana are consistent with the RHEME-initial configuration of the language. The elided-when-not-KNOWN nature of O's in the language is consistent with their obligatory (nominal) expression when they bear RHEME.

4.2.4. TOPIC as RHEME

The expression of TOPIC in Hixkaryana usually occurs sentence-finally, and, as Derbyshire notes (1985:156), is often right-dislocated. There are, however, occasions when TOPIC occurs sentence-initially, thus bearing RHEME. That is to say, there are occasions when "what-is-being-talked-about" is marked as "new" or "important" in some fashion relative to what has been said, or to what will be said.

There appear to be three occasions in which TOPIC occurs sentence-initially, bearing RHEME: first, to introduce a NEW participant as TOPIC; second, to re-instate a KNOWN participant as TOPIC, one who had been mentioned previously, but who had been "out of focus in the immediately preceding context" (Derbyshire 1979:72); third, to mark continued TOPIC. All three of these sentence-initial occurrences of TOPIC, within the nature of Hixkaryana, are consistent with the focal attention that is RHEME. First, then, as noted by Derbyshire (1979:71),

\[\text{Not all new information is placed into initial position, and is not, therefore, highlighted. New information highlighting occurs when it is going to be a major topic of the discourse ...}\]

For example, discourse-initial sentences often have the S "fronted", when it is going to be the "major topic of the discourse":

55) \text{kamara ti onono ryakoni}
The S, however, is not overwhelmingly "fronted" in discourse-initial sentences. Derbyshire (1979a:596) comments as follows on the statistics of the matter:

The statistical evidence with respect to discourse-initial sentences, based on a set of 30 published texts (Derbyshire 1965), is that: (i) a subject nominal occurs in the first sentences of 22 of these texts; and (ii) in 12 of them it is in final position, (O)VS, and in 10 of them it is in initial position, S(O)V.

Derbyshire continues that

It might have been expected that highlighting would always be a factor when introducing new participants at the beginning of a discourse. In fact, however, there are a number of ... [occasions on] which there is no need, or no desire, to highlight them ...

(Derbyshire 1977a:597)

The order that Derbyshire is here trying to explain is OVS discourse-initial sentences. One of the examples he gives of this order is:

57) yahutxho matkahekoni wosi
manioc-peel she-was-pounding-it woman
'A woman was pounding manioc peel'
(1977a:597)

If, however, one looks at the fuller context of this discourse-initial sentence, it becomes clear that it is not the S, but the rest of the sentence which is TOPIC in this case, and, consistent with what is suggested here, as a new TOPIC it is sentence-initial, bearing RHEME. The English gloss of the first seven sentences runs as follows:

1. The woman was pounding manioc peel. 2. There she was. 3. Action of pounding, she was doing it. 4. She was pounding it. 5. She was pounding manioc peel. 6.
There she was, working at it. 7. Action of pounding, she was doing it.

(1965:99)

The second occasion in which TOPICS occur sentence-initially is that in which a PARTICULAR which has earlier served as TOPIC is re-instated as such; obviously, on this occasion, the occurrence is discourse-medial. For example,

58) inyo ti nenyakon ha
    her-husband HSY he-was-seeing-it INTENSFR
    'Her husband was watching it'

(1979:72)

Derbyshire (1979:72) comments that in this sentence, the S (inyo 'her-husband') is "fron ted", and constitutes "the reinstatement of a topic (inyo) that had been introduced earlier, but had been out of focus in the immediately preceding context". Thus the content of RHEME in this case is not the introduction of a NEW particular, but rather the re-instatement of a KNOWN particular as TOPIC. (Note also that the re-instated TOPIC in this example is expressed nominally.)

The third occasion on which TOPICS occur initially is that in which continuation of TOPIC is marked. There are several ways in which this is accomplished, depending on the discourse status of the S: pronouns and word order are exploited in this regard. Initially, it may seem odd that continuation of TOPIC might require marking. However, when the RHEME-initial character of the language is taken into account, this becomes understandable. The semantics of RHEME are such that something "new" is expected at the beginning of each sentence. RHEME involves something that is REMOTE to focal awareness, and is characterized by a concentration on a DIFFUSE,
cognitively "active" semantics, rather than the FOCUSED, cognitively "stable"
semantics associated with TOPICS. Thus for the same TOPIC/Subject to be "in
focus" in two consecutive sentences is semantically odd, and requires overt
marking of the situation. Consider the following:

59) ito ti nehxakon ha kamara yohi ymo. noro ti
there HSY he-was INTENSFR jaguar chief-of AUG. he HSY

nonyetxkon ha . havana heno komo yonyetxkon. he-ate-them INTENSFR. visitor now-dead COLL he-ate-them.

'The jaguar chief was there. He used to eat them.
He used to eat the visitors.'

(1979:72)

Derbyshire cites these sentences as an example of "anaphoric focussing",
which, he says, "occurs when an anaphor is fronted for emphasis, where its
antecedent (or another anaphor of it) is in the immediately preceding sentence"
(Derbyshire 1979:72). Regarding (59), Derbyshire says that "noro ti has as its
antecedent kamara yohi ymo in the preceding sentence, and is fronted to keep
that participant in focus" (1979:72). It is suggested here that the "emphasis"
to which Derbyshire refers is that of continued TOPIC; since the "anaphor" of
noro '3rd person non-deictic' occurs in the immediately preceding context, its
discourse status is GIVEN, hence there is no need for nominal or pronominal
expression. However, with neither nominal or pronominal reference, RHEME
would fall on some other element, the verb perhaps. Sentence (58) above is an
example of re-instated TOPIC, whose prior mention is far enough away in the
discourse as to require nominal expression (akin to NEW discourse status).
Sentence (59) is analogous to (58) in that it is a re-instatement of TOPIC, but
its prior mention is very proximal, hence pronominal reference is used. Thus,
a pronoun in sentence-initial position has a discourse status of
RECOVERABLE; it bears RHEME because it is a TOPIC continued from the
immediately preceding sentence.

Derbyshire (1979:72; 1985:147) notes that there are "two contrasting positions of the third person pronoun (*noro* '3rd person non-deictic'; *nyamoro* '3rd person non-deictic, collective'); he says the choice between sentence-initial placement and sentence-final placement is determined by "whether or not the same 'topic' is in focus as in the preceding sentence" (1979:72). It is suggested here that the occurrence of this pronoun in final position has a discourse status of COMPUTABLE. For example:

60) krawame ihoko nehxatxkoni nyamoro with-difficulty occupied-with-it they-were they 'They were occupied with it with difficulty'

(1979:35)

In (60), the Complement is elided because GIVEN; the S is not expressed nominally, indicating it is KNOWN. However, the S also receives sentence-final, pronominal expression (*nyamoro*) this indicates that the S is KNOWN, but not GIVEN: its antecedent would be COMPUTABLE (mentioned earlier in the discourse). Thus it is not a continuation of TOPIC from the preceding sentence, which would require pronominal expression in sentence-initial position, as described above.

Derbyshire remarks (1985:147) that in addition to the use of these pronouns (*noro/nyamoro*) in sentence-initial position, involving "anaphoric focussing", that the particle *rma* 'SAME REFERENT, CONTINUITY' "also often occurs in a fronted phrase with this function". This particle will be further discussed below.

4.3.1. Hixkaryana: IMMEDIATE-REMOTE
This section presents specialized grammatical treatment that is accrued the expression of Direct Objects in Hixkaryana. Consistently, this treatment reflects a semantic REMOTEness.

4.3.2. Direct Objects as RHEME

It has been demonstrated that RHEME in Hixkaryana consistently occurs sentence-initially. In corroboration with Derbyshire’s accounts, when the O is nominally expressed, it most frequently occurs sentence-initially, resulting in the grammatical "basic" order of OVS. This sentence-initial placement of nominally expressed O’s thus consistently associates the O’s with the semantics of RHEME, which is consistently associated with sentence-initial position. Thus, when O’s are of NEW discourse status and require nominal (sentence-initial) expression, they are semantically REMOTE.

4.3.3.1. Definiteness and Indefiniteness: Identification of Particulars

As has been already mentioned, Derbyshire notes (1985:6) that the "categories" of definiteness and indefiniteness are not "morphologically marked" in Hixkaryana. Since these "categories" are necessary semantically, they must be communicated in a way other than by means of overt morphology.

It is significant in this regard that Derbyshire considers there only to be two ROLE relationships in the language: Subject (EXECUTOR) and Direct Object (EXPERIENCER). All nominals which do not function as S or O require the co-occurrence of what Derbyshire calls a "relator", which signals the manner in which non-ROLE nominals are to be interpreted (Derbyshire
Thus only S and O nominals do not have morphological indication of ROLE relationship. Grammatically, the indication of the status of a nominal as S or O is entwined with the expression of definiteness/indefiniteness, and with word order. Semantically, indication of the ROLES of EXECUTOR (S) and EXPERIENCER (O) is entwined both with indication of the discourse status of information (GIVEN-NEW) and with the expression of RHEME.

4.3.3.2. Prefixes

According to Derbyshire (1985:152),

The factor that most influences both the thematic and syntactic ordering of constituents in Hixkaryana (and in all other OVS and OSV languages for which I have seen data and/or descriptions) is the subject and object agreement in the verb affix system.

The significance which Derbyshire sees in connection with these prefixes he expresses as follows (1985:152):

[the prefixes allow] null realization of pronominal subjects and objects in many sentences in all types of discourse. The important consequence with regard to thematic patterning is that the unmarked theme or topic is often expressed only in the verb prefix.

These obligatory verbal prefixes code the persons of both the S and O (in portmanteau forms for transitive V's). For example,

61) koso w- ono- ye
deer 1st Subject/3rd Object- eat- dist. past compl.
'I ate the deer'

62) biryekomo y- otaha- no wosi
boy 3rd Subject/3rd Object- hit- imm. past woman
'The woman hit the boy'

A full paradigm of the verbal prefixes may be found in Table 1 (see Derbyshire...
1985:188-190 for discussion of their allomorphs).

When either the S or the O, or when both of them, are "recoverable" from the context, their respective nominals are obligatorily elided; however, the verbal prefixes always occur. Consider the following: (63a) is an OVS transitive sentence; in (63b), the S nominal is elided (resulting in an order of OV) with the S encoded in the verbal prefix; in (63c) the O nominal is elided (resulting in an order of VS) with the O encoded in the verbal prefix.

(63)a. toto y-ahosiye kamara
    man 3rd Subject/3rd Object, with preceding O-grabbed jaguar
    'The jaguar grabbed the man'
    (1985:32)

(63)b. toto y-ahosiye
    man 3rd Subject/3rd Object, with preceding O-grabbed
    'It (jaguar) grabbed the man'
    (1985:32)

(63)c. n-ahosiye kamara
    3rd Subject/3rd Object, without preceding O-grabbed jaguar
    'The jaguar grabbed him'
    (1985:32)

It is to be noted, both from the examples in (63), and from Table 1, that there is formal attention paid by the language as to whether or not a pre-verbal nominal is the O or not: the prefix for "3rd person Subject/3rd person Object" is dual in form, with y- occurring following an (expressed) O, and with ni- occurring "elsewhere". Since either S or O may be elided, and since an S may occur sentence-initially, there is, as Derbyshire notes, "potential ambiguity as to whether the NP is subject or object" (1985:32); this potential ambiguity, he says, "is usually avoided" by the choice of y- or ni-. The "potentially ambiguous" situation to which Derbyshire here refers concerns the case in which the identification of the S and the O are obvious from the discourse context, and both nominals are elided, as in
3rd Subject/3rd Object, without preceding Object - grabbed
(x) - grabbed -(x)
'It grabbed him' or 'He grabbed it'
(1985:32)

It may be suggested that (63d) is ambiguous only in isolation from the
discourse which occasioned it; hence the "potential ambiguity" which
Derbyshire suggests it is the purpose of this dual prefix to solve is, in the end,
a construct of the analyst himself.

Derbyshire (1985:152) describes the discourse status of elided S's and O's
as follows:

[such an elision] is always "known" or "given"
information in that it relates to either first or second
person referents, which are clearly shared information on
the part of the speaker and addressee, or third person referents, which are usually recoverable from the discourse
context, but may be shared information on other grounds,
in the sense that it is "activated" in the consciousness of
the addressee ...

Thus, it seems that information that is GIVEN is obligatorily elided. "First
and second person referents" are elided because such reference is "shared
information" between two interlocutors. "Third person referents" are not
necessarily "shared information" and thus are expressed nominally when they
are not thought to have GIVEN discourse status; if they are of GIVEN status,
they too are obligatorily elided.

Derbyshire does not differentiate at this point in his discussion between
the nominal expression or elision of S (EXECUTOR) and O (EXPERIENCER).
There is an important distinction to be made here, however. The choice
between the dual form of the prefix "3rd person Subject/3rd person Object" (y-
and ni-) is determined not by the discourse status of the S, but by that of the
O. That is to say, the language is more sensitive to the KNOWN/UNKNOWN (definite/indefinite) status of the O than it is to that of the S. This is consistently the case; it is only with third person objects that this sensitivity is morphologically reflected in the choice between two prefixes. With first and second person objects, such KNOWN/UNKNOWN status is obvious from "non-linguistic" context, as noted by Derbyshire (1985:152), quoted above. To recapitulate: 1) the prefix ni- '3rd Subject/3rd Object' occurs on the verb when there is no directly preceding O; there may or may not be a preceding ("fronted") S; that is, ni- occurs when the O is of GIVEN discourse status; 2) the prefix y- '3rd Subject/3rd Object' occurs on the verb when there is a directly preceding O; that is, it occurs when the nominal preceding the verb is the O, and of NEW discourse status. Whether the S is GIVEN (and its nominal elided) or NEW (and nominally expressed) is not morphologically marked in this way.

Davis (MS.b:139-141) notes that the discourse status of GIVEN, in which a Particular is KNOWN, IDENTIFIABLE, and CONSCIOUS, is one characterized by "an extreme of closeness to the here-and-now. Items ... are very immediate, salient, and within our awareness". The discourse status of NEW, in which a Particular is UNKNOW, UNIDENTIFIABLE, and NON-CONSCIOUS, is one characterized by being "remote and peripheral to our attention" (Davis MS.b:140). In Hixkaryana, the verbal prefixes under discussion are attuned to this IMMEDIATE (GIVEN) or REMOTE (NEW) semantics with regards to O's (not S's). The nominal expression of O's is without fail the introduction of a NEW participant; hence, O's in the language are consistently REMOTE.
There are two other occasions in which prefixes function in the language. It is suspected that they serve to indicate the semantics of IMMEDIATE-REMOTE in a consistent manner, but more study is necessary to be certain of this. These occasions are mentioned here, but are beyond the scope of this chapter to investigate more fully.

The first involves the marking of possession on Nouns. Verbs have as prefixes only the person-marking ones just described; as well they have suffixes which generally mark tense and aspect, and sometimes mood and number. In the nominal morphology, prefixes on the Noun mark person of the possessor, and suffixes mark possession (N.B.), tense, and number. What is interesting for the present discussion is that there is overlap in form in many of the "nominal" and "verbal" prefixes. Derbyshire (1979:98) notes this also, and states the similarity as an observation on occurrences:

Comparing the list of noun prefixes with the paradigm for person-marking prefixes in the verb ..., it can be seen that there is exact correspondence between the possessives and the group of verb prefixes which mark the direct object occurring when the subject is third person, except for III, where the basic allomorph in the noun prefix is i- and that in the verb is ni- ...

It is proposed here that this overlap is not accidental, and that it reflects the encoding of a similar semantics, regardless of "Noun" and "Verb".

As has been said, the "verbal" prefix y- occurs when there is an overtly expressed nominal preceding the Verb, and it is to be interpreted as an O of NEW discourse status. There is also a "nominal" 3rd person prefix y- (with phonologically conditioned allomorph θ-), which occurs only when the immediately preceding constituent is a nominal, and in fact the person possessing:
64) waraka y- owa- ni
Waraka 3rd-person chest possessed-item
'Waraka's chest'
(1985:200)

65) waraka 0- kanawa- ri
Waraka 3rd-person canoe possessed-item
'Waraka's canoe'
(1985:200)

If there is no preceding nominal in a possessive construction, the prefix is i-
(phonologically conditioned allomorphs u-, stem-initial change o- to e-, θ-):

66) e- wa- ni
3rd-person chest possessed-item
'his chest'
(1985:200)

67) i- kanawa- ri
3rd-person canoe possessed-item
'his canoe'
(1985:200)

Table 2 summarizes the situation concerning the form, occurrence, and
syntactic function of these four (?) prefixes. Semantically, there is coherence in
the quasi-overlapping obvious in Table 2. In the case of "nominal" i- (as in
[66] and [67]) and "verbal" ni- (as in [63c] and [63d]), the possessor and the
O’s (respectively) are inferrable from the discourse context (they are KNOWN
and IDENTIFIABLE), and have thus been elided. However, in the case of both
"nominal" y- (as in [64] and [65]) and "verbal" y- (as in [63a] and [63b]), the
possessor and O’s (respectively) are UNKNOWN and require nominal
expression.

These prefixes constitute part of the manner in which Hixkaryana
expresses grammatical definiteness/indefiniteness, or the semantics of
GIVEN-NEW discourse status. There is formal attention paid by the language
to whether or not the preceding nominal constituent is KNOWN (and therefore elided, and occurring only in the verbal/nominal prefix) or UNKNOWN (and therefore requiring nominal expression, as well as the prefixal element). In the case of Noun-like stems, this attention coincides with the person of the possessor, and in the case of Verb-like stems, it coincides with the O.

The second occasion in which prefixes occur is in the formation of subordinate, "nonfinite" clauses. Throughout this chapter, discussion centers on the main clause formation and the semantic motivation for its word order. Thus this point is merely mentioned here, and is a direction in need of further study. Derbyshire (1979:23) says that the "dominant form of subordination" in Hixkaryana involves "nominals derived from .... verbs". In these derived nominal "pseudo-clauses",

the direct object of an underlying transitive, and the subject of an underlying intransitive, are marked in the same way as possessors in noun phrases (the person-marking prefixes in verbs ... and nouns ... overlap, but there are some formal differences, and at this point the derived nominals follow noun forms rather than the verb).

The Subject of an "underlying transitive", however,

surfaces as a postpositional phrase with the noun followed by wya 'to, by', which is normally an indirect object marker. (1979:23)

For example:

68) a. oxhe nhananihyatxhe wosi
good she-teaches-them woman

(3rd S/3rd O-teach -NONPAST COLL )

'The woman teaches them well'

(1979:24)

b. oxhe ihananihniri komo wosi wya
good teaching-of-them COLL woman by
(i  hananihi  -ri )
(3rd p. -teach  -ACT.NOMLZN.-POSSN)
'The teaching of them well by the woman'

The understanding of the nominal/verbal prefixes which has been presented here suggests that in their occurrence in main clauses they reflect the KNOWN or UNKNOWN discourse status of the preceding nominal. That this GIVEN-NEW status is an instantiation of the semantics of IMMEDIATE-REMOTE has also been pointed out. It is suspected, therefore, that the prefixes in these nonfinite constructions also will reflect the semantics of IMMEDIATE-REMOTE; in this case, it appears that O's and intransitive S's may be more "prototypically" REMOTE, while transitive S's may be more "prototypically" IMMEDIATE.
<table>
<thead>
<tr>
<th>S</th>
<th>1st + 2nd p.</th>
<th>1st p.</th>
<th>2nd p.</th>
<th>3rd p.</th>
<th>(S only)</th>
<th>(S only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st + 2nd p.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st p.</td>
<td></td>
<td>ki-</td>
<td></td>
<td>i-</td>
<td>ki-</td>
<td>w-</td>
</tr>
<tr>
<td>1st + 3rd p.</td>
<td></td>
<td>o-</td>
<td></td>
<td>ni-</td>
<td>ni-</td>
<td>n-</td>
</tr>
<tr>
<td>3rd p.</td>
<td>ki-</td>
<td>o-</td>
<td>ro-</td>
<td>(−O) ni-</td>
<td>ni-</td>
<td>n-</td>
</tr>
<tr>
<td>2nd p.</td>
<td></td>
<td>mi-</td>
<td></td>
<td>mi-</td>
<td>mi-, o-, ow-</td>
<td>man-, m-</td>
</tr>
</tbody>
</table>

Table 1: Full paradigm of verbal prefixes  
(from Derbyshire 1985:188)

<table>
<thead>
<tr>
<th>FORM</th>
<th>SYNTACTIC FUNCTION</th>
<th>OCCURRENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERBAL</td>
<td>y-</td>
<td>¬marks 3rd person S, 3rd person O</td>
</tr>
<tr>
<td>NOMINAL</td>
<td>y-</td>
<td>¬marks 3rd person possessor</td>
</tr>
<tr>
<td>VERBAL</td>
<td>ni-</td>
<td>¬marks 3rd person S, 3rd person O</td>
</tr>
<tr>
<td>NOMINAL</td>
<td>i-</td>
<td>¬marks 3rd person possessor</td>
</tr>
</tbody>
</table>

Table 2: Prefixes
4.3.3.3. Particle *rma* 'SAME REFERENT, CONTINUITY'

In the discussion of TOPICS in sentence-initial position, it was remarked that Derbyshire (1985:147) says that the particle *rma* can function in a manner similar to that of third person pronouns in sentence-initial position, to effect what Derbyshire calls "anaphoric focussing". It was concluded that the pronouns in this position mark continuity of TOPIC from the immediately preceding sentence.

The particle *rma* has differences and similarities to the semantics of third person pronouns. The pronouns (obviously) always have third person referents as their antecedents; the particle, however, according to Derbyshire (1979:58) "may refer to a person or thing, or to an action". His gloss for *rma* in his earliest publication (1965:9) also indicates that the particle does not necessarily refer to a particular in discourse: "'ainda', inica continuacao de contexto" ['same'; indicates continuation of context', EAC]; he notes (1979:58) that its "main function is anaphoric".

Examples of *rma* apparently referring to particulars are as follows:

69) *ito rma oknomkewo mokyamo nehxatxkon hati*
   *there SAME ones-who-were-left those-people they-were HSY*
   'They were the people who were left there'
   (1985:53)

70) *towahke rma totokom hok nehxakoni karaywa*
   *friendly CONTINUITY people occupied-with he-was non-Indian*
   'Yet the non-Indian was friendly towards the people'
   (1985:75)

An example of *rma* apparently having a verbal anaphor is as follows:

71) *iro rma narymaye , masku rma that-thing SAME-REF it-throw-him , dysentery SAME-REF*
   'The dysentery (I've been talking about) killed him'
   (1979:58)
Derbyshire comments that in (71), "rma occurs twice modifying words whose antecedent occurs fourteen sentences earlier in the discourse, the verb masuhtaxe 'he had dysentery' (1979:58). Another example is found in (72), concerning which Derbyshire says, "rma modifies a word whose antecedent is a synonym, amamhehra 'not delaying', which occurs five sentences earlier":

72) nomokye harha ti iramampira rma
   he-came back HSY not-delaying SANE-REF
   'He came back, still not delaying'

(1979:58)

(Both iramampira 'not delaying', and amamhehra 'not delaying' appear to be nominals derived from verbs; the scope of rma is again apparent.)

What is significant for the present discussion is that apparently rma has only the S or the V (both possible TOPICs) as its "antecedent", but not the O. An examination of forty-nine sentences in which rma occurred did not result in any examples in which the particle had the O as its antecedent. The previous discussion in this chapter of occurrences of TOPIC as RHEME brought up the role of resumptive pronouns in the language. It is to be now noted that these pronouns consistently refer to the S, but not to the O. It appears that while both S's and V's may function as TOPICs, that O's cannot function in this way. Semantically, the inherent REMOTEness associated with O's in Hixkaryana is incompatible with the IMMEDIATEness characteristic of TOPICs. Thus, grammatical O's are consistently associated with the semantics of RHEME in Hixkaryana. Since sentence-initial position has been demonstrated to be consistently exploited for the expression of RHEME in the language, the O occurs sentence-initially.
4.4. Concluding Remarks on Hixkaryana

This chapter has provided an interpretation of Hixkaryana word order which explains the semantics of its "basic" OVS sentential syntax. It was first suggested that the language is consistently RHEME-initial. Davis has noted that there are other languages with this functional order, and examines Bella Coola (MS.b:68ff.) and Tagalog (MS.b:216ff.) in this light; Hoff (1978) notes the same order for another South American language, Carib. Thus, there are other languages which are RHEME-initial.

The second half of this interpretation involved the examination of specialized syntax involved with Direct Objects in Hixkaryana. It was suggested the Direct Objects are consistently associated in the language with a cognitively active, "REMOTE", or "DIFFUSE" semantics that is also associated with RHEME. Thus Direct Objects, when nominally expressed, occur most often in that position in Hixkaryana which is reserved for RHEME's expression: sentence-initial position. Davis has noted that there are other languages in which Direct Objects are consistently associated with RHEME: Haya and Zulu (MS.b:45ff.).

Thus, the two semantic configurations discussed in this chapter for Hixkaryana are not unique in themselves; what is unusual cross-linguistically is for these two configurations to occur in the same language.

This, then, is a story of Hixkaryana word order. Another OS language, Tzotzil (VOS) will be discussed in the next chapter. In the story of Tzotzil word order, certain themes which have emerged in the story of Hixkaryana will reappear.
CHAPTER FIVE

Tzotzil (Mayan): VOS

5.1. Extent of VOS Word Order in Mayan

Within the Mayan family, there is a concentration of languages which have been classified as VOS in basic word order. This chapter will examine one of these languages in some detail: Tzotzil. Reference will be made to other Mayan VOS languages at relevant points. An explanation of the claimed VOS basic word order of Tzotzil will be presented, in a manner similar to that utilized in the preceding chapter for Hixkaryana.

In Pullum's catalog (1981:151) of Object-Subject languages, there are thirty languages listed. Of these, thirteen are said to be VOS, and another five as "VOS?". This gives a total of eighteen VOS languages; what is to be noted is that eleven of these eighteen are of the Mayan family. That is to say, over one-third of the OS languages in Pullum's list are Mayan and VOS.

Quizar (1979), in her comparative study of Mayan word order, arrives at a similar count. She lists twenty-two Mayan languages in all; eight of these she says are VOS, and another four as "VOS?" (1979:93-5), resulting in a total of twelve VOS languages in Mayan, or more than half of the total twenty-two. These counts are presented here to indicate that OS word order appears to be widespread in Mayan. Though the main discussion here will be concerned with Tzotzil, some discussion of other Mayan VOS languages will occasionally be brought in. This will both strengthen the semantic interpretation being given here, and intimate that perhaps this semantics is not specific to Tzotzil, but
has a wider applicability with the Mayan VOS languages. Thus, though each language will idiosyncratically employ the limited number of elements of discourse structuring mentioned in Chapter Three, it is suggested that perhaps a similar semantics is observable in OS languages; it is a semantics which has as one of its syntactic effects that Objects precede Subjects in what has been called "basic word order".

5.2. Tzotzil (VOS) and Word Order Typology

Tzotzil is spoken in the state of Chiapas in southern Mexico. Laughlin (1977:5-6) writes that it is spoken by over 120,000 speakers, ranking seventh highest in the number of speakers of Mayan languages. The dialect which will be presented in this chapter is that of Zinacantan, spoken (according to Laughlin) by 12,000 Zinacantecs, in nineteen townships in Chiapas. This dialect is that studied by Laughlin, Haviland, and Aissen; the publications of Haviland and Aissen form my primary sources.

The following discussion, then, will be guided primarily by a consideration of the language Tzotzil. Recall Greenberg's (1963) statement that of the six logically possible orders SVO, SOV, VSO, VOS, OSV, and OVS, only the first three "normally occur as dominant orders" (Greenberg 1963:76). Languages in which the O preceded the S in basic word order (that is, VOS, OSV, and OVS), Greenberg claimed "do not occur at all, or at least are excessively rare".

Subsequent discussion of this statement was reviewed at the beginning of Chapter Four, and will be returned to in Chapter Six. Chapter Four discussed a language of one of these rare, or exceptional, word order types: Hixkaryana (OVS). The present chapter will discuss Tzotzil (VOS) which is another of
these rare word order types. The third OS order, OSV, will not be discussed in
the present work. However, it is to be noted that such languages do exist.
Pullum (1981:151), for example, in his catalog of thirty OS languages lists
three as OSV and one as "OSV?". These languages are discussed in

Tzotzil, and Mexican-Guatemalan languages in general, have received
much more attention from linguists and linguistic anthropologists than have
the Amazonian languages such as Hixkaryana; this is due, doubtless, to the
physical problems of accessing speakers of the latter. This relatively greater
attention has resulted in a larger literature available on the Mayan languages.
As well, both the longer length of time and the larger number of researchers
involved in Mayan studies has resulted in a more explicit discussion in the
literature of the implications of Mayan languages for word order typology.

In the literature on Tzotzil of which I am aware, the language is simply
stated as being VOS in basic word order (Aissen 1987:1, 2; Aissen 1983:275;
am not aware of a detailed discussion of how this order was determined to be
"basic" in the sense intended by word order typologists.

Though, by comparison with the Amazonian language Hixkaryana there is
more published material on Tzotzil language and culture, not all of it has
proved usable for the present chapter, and for varying reasons. There has been
much interesting work published by linguistic anthropologists (e.g., Bricker
1977, 1980). The purpose of these publications is different from that of the
present exercise, and the language data available in them do not carry literal
glosses; it is both beyond the scope of the present work and beyond the writer’s
knowledge of the language to attempt to provide such glosses. My main source
of linguistically glossed data is Aissen, who has extracted much from Laughlin,
and has provided literal glosses for the examples she uses. Haviland (1981) is
a grammar of Tzotzil written in Spanish, from a non-technical stance;
examples from this source have no literal glosses in the original. I have
provided such glosses, gleaning them from Haviland’s work, and translated all
excerpts from Haviland into English. Sources are acknowledged with each
linguistic example; literal glosses appear as in my sources, thus resulting in
some minor discrepancies. These, however, should not prove detrimental to my
discussion.

5.3. General Remarks on Tzotzil

Tzotzil has been classified as VOS, and this is the word order which occurs
in unmarked transitive sentences with nominal S’s and O’s:

73) 7ī- s- poxta xun 1i j7īl0l- e
cp A3 care Xun the shaman cl
'The shaman treated Xun'

(Aissen 1987:90)

Other word orders occur frequently as well, and these will be introduced durin.g
the course of this chapter.

Aissen (1987:1) gives the following introductory remarks on the language:

Subjects and objects are not marked for case, but more
peripheral relations are. The predicate agrees in person,
and sometimes in number, with its subject and direct
object. The agreement system is ergative. Non-emphatic
pronouns do not occur in surface structure.

The ergative agreement system means that transitive O’s and intransitive S’s
bear the same set of affixes; this combination is known as the absolutive relation, and the affixes which mark it are termed by Mayanists as "Set B affixes". Transitive S's are marked by another set of affixes; this constitutes the ergative relation, and Mayanists refer to the corresponding set of affixes as "Set A". In examples from Aissen, literal glosses of these affixes indicate either a Set A or Set B affix, and the number of the person (Aissen 1987:2).

Set A affixes also, Aissen (1987:4) writes, mark person of the possessor on the possessed item. Haviland (1981:290-281) notes the variety of uses of constructions marked by these Set A affixes, saying that "In each example, the relation between possessor and possessed is different". These affixes in Tzotzil are prefixes, though "Tzotzil also has a set of cognate suffixes that can be used in place of or in addition to the Tzotzil personal prefixes as objects of transitive verbs" (Hopkins 1977:199). Bricker (1977), in her comparative study of these Mayan affixes of "pronominal inflection" assigns Tzotzil an "anomalous position", in this regard; she writes:

Both the O-S-V and S-V-O patterns of pronominal inflection occur in Tzotzil. The Tzotzil dialects spoken in Zinacantan, Chamula, San Andres Larrainzar, San Pedro Chenalho, San Pablo Chalchihuitan, and Huitan have three pronominal series .... The first two series function like the A and B pronouns in the O-S-V languages .... The third series in Tzotzil functions like the set B pronouns in the S-V-O languages ...

She further notes (1977:9) that

In Zinateco Tzotzil (as well as the dialects of Chamula, San Andres Larrainzar, San Pedro Chenalho, and San Pablo Chalchihuitan), the O-S-V pattern is associated with incomplete and complete temporal-aspectual inflection, and the S-V-O pattern occurs with perfective and subordinate constructions and with stative verbs ...

Tzotzil has a sister language, Tzeltal; Hopkins (1977:198) writes that
Together, Tzeltal and Tzotzil constitute a single linguistic area. Speakers of both languages inhabit the same geographic area, share a common culture, and appear to consider Tzeltals and Tzotzils as a single kind of Indians, distinct from Chols, Tojolabals, Zoques, Zapotees and other nearby Indian groups of whose existence they are aware. Any speaker of Tzeltal or Tzotzil will indiscriminately refer to the speech of any other Tzeltal or Tzotzil community with the same term: bac'il k'op 'true (proper) speech'.

The principle morphological feature which can serve as an "unambiguous diagnostic" between Tzotzil and Tzeltal dialects is that the Tzotzil affixes are prefixes, while those in Tzeltal are suffixes (Hopkins 1977:199).

The significance of the opposition between Set A (ergative and possessive) and Set B (absolutive) affixes combined with the lack of case-marking on S's and O's has drawn the interest of linguists Aissen (1983, 1984b, 1987) for Tzotzil, and Walter (1980) for Tzeltal. Aissen, working within the theoretical framework of relational and arc pair grammar, discusses the significance of these syntactic configurations in terms of "Indirect Object Advancement" and "Possessor Ascension". As noted by Walter (1980:244),

Aissen's analysis depends upon the claim made by relational grammar that the syntactic relations -- subject, direct object, indirect object -- are universal syntactic primes.

This fact is also stated by Aissen (1983:273). Walter provides a discussion more in the spirit of that guiding approach in the present work; he concludes his own discussion as follows (1980:247):

The distinctiveness of ergativity has served as a partial motivation for my thesis that the fundamental linguistic opposition in Tzeltal is between affector and affected.

The discussion which follows, however, will not have as its prime focus these issues of ergativity and case-marking; their relevance is beyond the scope
of this chapter. Rather the interest will be the semantic function of sentential position in Tzotzil, and what causes O's to precede S's in "basic word order".

5.4. Tzotzil: DIFFUSE-initial

5.4.1. Reports in the Literature

It has been noted by a number of researchers that Tzotzil, and a number of Mayan languages, are functionally "comment-topic" in sentential order. Haviland (1981:19) states, for example, that

A simple sentence in Tzotzil has the basic order

Comment     Topic.

That is to say, the topic, or the thing being talked about, follows the comment, or that which the sentence says about the topic.

Pickett, in her study of Mexican languages with regards to Greenberg's (1963) universals, "checked sixty-two languages" (1983:530) with the help of her students and colleagues. She remarks on "Hockett's statement that there is a general tendency [in all languages] for comment to follow topic"; she concludes from her study, however, that "we find that the Mexican languages do not confirm this generalization" (1983:536). Laughlin (1977:6) writes that "Tzotzil has much in common with its Guatemalan [Mayan] relative, Quiche." He then quotes Edmonson (1971:xii) describing Quiche discourse, and says that this also applies to discourse in Tzotzil:

the sentence "leans forward .... In something of the sense that German saves verbs for a final triumphant closure to a thought, Quiche saves nouns. There are markers to show that the noun is coming. Its number, for example, may be revealed early, and other attributes may be thrown in along the way, all contributing to the forward movement toward the person or object who would be in English the antecedent!"
Recall the character of RHEME being employed in this study, described in
Chapter Three and implemented for Hixkaryana in Chapter Four; Edmonson's
description of the feeling of Quiche discourse is easily compatible with one
structured by sentences which begin with RHEME-like elements. Edmonson's
description also reminds one of what Derbyshire called "cataphoric reference"
with regards to the identification of participants in Hixkaryana.

5.4.2. Expressions of Emphasis

Mayan languages are not strictly verb-initial; there is a strong tendency
for a sentential element which is emphasized to occur sentence-initially (e.g.,
Tojolabal Maya [Brody 1984a]; K'e'ch'i [Pinkerton 1978]; Quiche [Quizar
1979:81-83]; Jacaltec [Quizar 1979:77-80]).

Haviland (1981:24) provides examples of this in Tzotzil with regards to
locative and temporal expressions. Sentence (74) involves a locative expression
without emphasis; in sentence (75) the location is emphasized, and occurs
sentence-initially:

74) 7oy ch'ivit ta Jobel
     there—is market definite preposition San Cristobal
     (hereafter "def-prep")
     'En San Cristobal hay mercado'
     'There is a market in San Cristobal'
     (Haviland 1981:24)

75) ta Jobel 7oy ch'ivit
def-prep San Cristobal there—is market
     'Es en San Cristobal que hay mercado'
     'It's in San Cristobal that there is a market'
     (Haviland 1981:24)

An emphasized temporal expression, analogous to (75), also results in a
cleft-sentence in English, indicating the sentence-initial position of RHEME.

76) ta 7ak'ubal 7oy pukuj.
def-prep night there-are devil
ta 7ol k'ak'al ch'abal
def-prep heavy day there-aren't

'Es por la noche que hay diablos; a mediodía no hay'
'It's in the night that there are devils;
(it's) at midday (that)there aren't any'

(Haviland 1981:24)

Aissen (1987) also gives examples indicating the emphatic use of
sentence-initial position for what she calls "adverbials" ([77], [78], [79]) and
"predicational elements" ([80], [81]):

77) k'ajom ta x- vay
    just icp sleep
    'He just slept'

(Aissen 1987:18, from Laughlin 1977:55)

78) te 7i- bat ta Saklum
    there cp go to Saklum
    'He went there to Saklum'

(Aissen 1987:18)

79) jal te- on
    long there Blsg
    'I was there a long time'

(Aissen 1987:18, from Laughlin 1977:203)

80) chopol xa kom
    bad cl remain
    'He was left in bad shape'

(Aissen 1987:18, from Laughlin 1977:85)

81) te vayem 7i- s- ta- ik
    there asleep cp A3 find pl
    'They found him asleep there'

(Aissen 1987:18, from Laughlin 1977:143)

In an unemphatic context, the preverbal elements in (77) through (81) would
occur postverbally. Sentence (82) provides another example of this; the context
given here again indicates the emphasis expressed in Tzotzil by sentence-initial
position:

82) There were no cars then, so --
There is a particle in Tzotzil, *ja7*, which Aissen simply glosses as 'i'; she says (1987:19) that this particle "is emphatic, drawing attention to some element in the clause which follows". It is interesting to note that an identical particle *ja7* occurs in Tojolabal Maya, and that Brody, in her thorough study of discourse function in that language, calls it a "cleft-word". She writes of Tojolabal *ja7*:

In transitive sentences, the cleft-word always is inflected for third person, and always appears at the beginning of a clause, though a constituent may be focused or topitized and hence appear pre-verbally. The focus of the cleft is the material which follows the cleft-word .... All of the sentence following the *ja7* is focus-of-cleft. For those sentences where *ja7* is the first word, the whole sentence is the focus-of-cleft, and the presupposition is to be found in surrounding discourse.

Jacaltec has an analogous particle *ha7* which also appears to occur sentence-initially, and which is known as a "focus marker" (Quizar 1979:79). Though the examples of Tzotzil *ja7* which are available to me ([83] through [86]) are not glossed with clefted constructions in English, it is very possible that they could carry such emphasis as Brody has found in Tojolabal. Thus, it may be that this particle serves to indicate an exclamatory feeling for the whole sentence, rather than just for "some [one] element", since the latter is accomplished in Tzotzil by sentence-initial positioning. However, what is most relevant in the present discussion is to note that, in the examples I have, this emphatic particle occurs sentence-initially:

83) *ja7 te tzinil taj garafon 7une*
    1 there crowded that jug cls
    'Those jugs were packed in tight there'
84) ja7 j- chi7uk li Xun- e
   ! Al with the Xun cl
   'I'm with Xun'

85) ja7 j- chi7uk lok'el 7un
   ! Al with leaving clis
   'I left with them'

86) ja7 no la te lok' ...
   ! cl cl there left
   'She just went outside there ...'

A final example will complete this section; the emphatic nature of sentence-initial position is again apparent:

87) batz'i 7ep yayijem- on yu7un ni vinike
    very lots wounded- B1 by the man
    'I'm terribly wounded by the man'

5.4.3. Numerical and Quantitative Expressions: "Discontinuous Constructions"

Edmonson (1971:xii), quoted above, described Quiche discourse as "saving nouns" till the end of a sentence, though there were "markers to show that the noun is coming". He noted specifically that "[the noun's] number ... may be revealed early", while the noun to which it referred occurred late in the sentence. A similar situation appears to hold for Tzotzil; this section examines instances in which either a specific numeral or an indefinite plural (7ep 'lots' or 7epal 'many') modify either a Subject or an Object. The differences in sentential position which may be observed in these expressions will contribute in several ways to the present discussion. First, they will add evidence to the
suggestion that sentence-initial position is reserved for the expression of a
DIFFUSE, RHEME-like semantics. Second, they will add evidence to the
suggestion that sentence-final position indicates a FOCUSED, TOPIC-like
semantics (this suggestion will be discussed in detail below). Third, it will
reveal a specialized treatment of O's in Tzotzil, which will aid in an
understanding of their nature in the language.

Some introductory remarks must be made concerning numerical
expressions in Tzotzil. There are numbers in the language (e.g., cha7 'two', 7ox
'three', chan 'four') to which are added suffixes known as "numerical
classifiers". Aissen writes that

The classifier restricts the class of items being counted,
usually in terms of some salient physical property ... vo7 is
used for counting humans, kot for quadrupeds, pej for
squattish things, and so on. (1987:6)

(Berlin [1968] represents an exhaustive study of these numerical classifiers in
Tzeltal.) The numerical classifiers are not suffixed to the indefinite plural
expressions 7ep 'lots' and 7epal 'many'.

There are quantitative sentences in Tzotzil which reflect the basic VOS
order in the language; in these sentences, the quantitative expression occurs
preceding its nominal (either the S or the O), with the latter occurring in their
"normal" positions. In these constructions, the number may represent NEW
information, in which case functionally they may be examples of "unmarked
RHEME".

Sentences (88) through (92) exemplify this "normal" positioning of a
Subject and its quantifier:

88) 7i- cham cha7- kot j'- chitom
cp die two nc A1 pig
'Two of my pigs died'

(Aissen 1987:261)

89) 7i- vaychin i 7ox- vo7 j7iloletike
cp dream the three nc shamans
'The three shamans dreamt'

(Aissen 1987:6, from Laughlin 1977:95)

90) ba s- man- ik tal kaxlan vaj 7epal kremotik
went A3 buy pl here bread many boys
'Many boys went to buy bread'

(Aissen 1987:269)

91) 7i- s- k'el- ik k'in 7epal tzebetik
cp A3 watch pl fiesta many girls
'Many girls watched the fiesta'

(Aissen 1987:269)

92) 7i- s- k'el- ik k'in cha7- vo7 viniketik
cp A3 watch pl fiesta two nc men
'Two men watched the fiesta'

(Aissen 1987:264)

Note the expected, sentence-final occurrence of the S and its quantifier in (88)
through (92). Sentences (93) through (96) exemplify Objects with their
quantifiers; note that in each example, the quantifier and the nominal it
modifies occur in the O's "normal", second-to-last position:

93) 7i- y- elk'an cha7- kot chij
cp A3 steal two nc sheep
'He stole two sheep'

(Aissen 1987:261)

94) 7i- s- sa7 7ep s- kriara- ik
cp A3 seek lots A3 maid pl
'They acquired many maids'

(Aissen 1987:261, from Laughlin 1977:73)

95) 7i- y- ich' 7ep tak'in li viniketik- e
cp A3 get much money the men cl
'The men received a lot of money'

(Aissen 1987:262)

96) 7i- s- man- ik 7ep kaxlan vaj li kremotik- e
cp A3 buy pl lots bread the boys cl
'The boys bought lots of bread'

(Aissen 1987:263)
It is construed, in sentences (93) through (96), that no special emphasis is being accrued either the quantity alone, or the quantity with its noun. These examples will be seen to contrast with those that follow, in which such emphasis is indeed present.

It is possible to have the quantity occur preverbally (that is, in sentence-initial position), while its (S or O) nominal occurs in its "normal" position. Sentences (97) through (100) are examples of this preverbal quantifier modifying the (sentence-final) S:

97) 7ep 7i- 1aj ti Pinedae
    lots cp die the Pinedists
    'Many of the Pinedists died'

(Aissen 1987:253)

98) 7ep xa ch'ay y- osi l ti krixchanoetik lej- e
    lots cl lost A3 land the people there cl
    'A lot of the people there's land was lost'


99) 7ep 7i- s- nak' s- ba- ik ta ch'en li viniketik- e
    lots cp A3 hide A3 self 3pl in cave the men cl
    'Lots of men hid themselves in the cave'

(Aissen 1987:257)

100) 7ep 7i- s- tzob s- ba li sapo- e
    lots cp A3 gather A3 self the toad cl
    'Lots of toads got together'

(Aissen 1987:257)

Examples of a preverbal (sentence-initial) quantifier modifying an O which occurs in its "normal" second-to-last position are provided by sentences (101) through (106):

101) 7ep 7i- s- mal- be s- ba kalto li Xun- e
    lots cp A3 spill io A3 self soup the Xun cl
    'Xun spilled a lot of soup on himself'

(Aissen 1987:258)

102) 7ep 7i- s- vok'- be s- ba- ik y- osi l li viniketik- e
    lots cp A3 break io A3 self pl A3 land the men cl
    'The men hoed a lot of each other's land'
(Aissen 1987:258)

103) 7ep 7i- s- k'el- ik k'in li tzebetik- e
lots cp A3 look pl fiesta the girls cl
'The girls saw many fiestas'

(Aissen 1987:255)

104) 7ep ta- s- jim- ik bala li solterotik- e
lots icp A3 fire pl bullet the soldiers
'The soldiers fired many bullets'

(Aissen 1987:255)

105) cha7- vo7 7i- s- kolta tzebetik li jsoktometik- e
two nc cp A3 help girls the Chiapanecos cl
'The Chiapanecos helped two girls'

(Aissen 1987:260)

106) 7ox- vo7 7i- s- mil- ik viniketik li jsoktometik- e
three nc cp A3 kill pl men the Chiapanecos cl
'The Chiapanecos killed three men'

(Aissen 1987:260)

It is important to note that in sentences (97) through (106), in which a quantifier has been "fronted", it is unambiguous to a Tzotzil speaker as to which nominal the quantifier refers. Apparently it is the discourse and "real-world" context which causes there to be no ambiguity. Aissen (1987:258ff.) attempts an explanation for the lack of ambiguity in these sentences by means of a complicated syntactic rule:

Tzotzil Initial Absolutive Bindee Rule (informal): If a quantifier q in a Tzotzil PN S binds a nominal b which does not head an initial absolutive arc in clause c, then there is no PN R which is identical to S except that q binds a nominal b which heads an initial absolutive arc in d, where d in R is the correspondent of c in S. (Aissen 1987:261)

However, it is simpler and probably closer to the facts to say that ambiguity is nonexistent because sentences occur in speech situations! This suggestion is corroborated by an incident recounted by Aissen herself, concerning "speaker reaction" to the following sentence:  .
Aissen writes:

One interpretation is indicated in the translation. But one informant, laughing, translated it, 'I gave three pieces of candy to the boys'. He explained that the candies had to be shaped like humans. Noun stems are not arbitrarily assigned to classifier sets; rather a classifier imputes properties to the referent of the relevant nominal .... Of course, speakers will vary in the interpretations they 'see', both among themselves and from occasion to occasion.

The final alternate order of quantifiers and their nominals occurs in the situation in which both the quantity and its nominal are "fronted", and occur sentence-initially. In examples (108) through (112), note that the English gloss reflects the fact that these fronted constituents bear RHEME.

(Aissen 1987:6)


(Aissen 1987:269)

(Aissen 1987:269)
(lit.: 'Much was the money that he was given')

(Aissen 1987:262)

What is most significant is that (108) through (112) consistently have S's and their quantifiers occurring in sentence-initial position. I have not found any examples which reflect the analogous situation with O's. That is, I have no examples in which both an O and a quantifier modifying it occur in the sentence-initial emphatic position.

5.5. Tzotzil: FOCUSED-final

5.5.1. "Unmarked" and "Marked" Topic and Focus

Researchers on Mayan discourse have explored the effects and interplay of word order with morphosyntax, and have determined that what I have called "discourse motivation" (see Chapter Three) is much more relevant to an understanding of the workings of these languages than it is for "grammaticalized word order languages" such as English. For example, Brody (1984a) examines Tojolabal Maya in this regard, and Pinkerton (1976b, 1978) provides a study of K'ekchi discourse. The suggestions made in the present work for Tzotzil are not meant to be as exhaustive as these detailed studies, as the purpose here is somewhat different; the comparison is made to indicate that similar discourse forces may be working in Tzotzil.

Some preliminary remarks on this matter for Tzotzil are made by Myhill (1983). He provides little language data in his paper, but rather relies on statistical data; his comments on Tzotzil, however, will provide useful hints towards an understanding of word order in the language. Myhill, proposing a typology of VO languages, suggests that discourse in such languages is
organized in terms of "unmarked" focus and topic constructions, and "marked" focus and topic constructions. His observations will be seen to agree with the present suggestion that Tzotzil is DIFFUSE-initial and FOCUSED-final.

Calling Tzotzil "solidly VOS" (1983:157), Myhill notes that "marked topics are preverbal, while subjects of type B are not". "Type B subjects" he describes as occurring in those sentences in which

the predicate indicates a change in the information status of the subject rather than any other character, context, or set in the discourse register. (1983:156)

This indicates that in sentences of "normal" VOS order, the verb bears RHEME, in that it is "the predicate" that "indicates a change". In sentences which are Subject-initial, the S, though TOPIC, bears "marked" RHEME. The suggestion that the Tzotzil verb bears RHEME in VOS order is corroborated by Myhill's statement that, in Tzotzil, "sentential focus and verb phrase focus ... [are] associated with the same position" (1983:161). Myhill further notes (1983:157) that both S's and O's may be "topicalized": "subjects and objects may be topicalized by being put in initial position and being marked with the enclitic -e". This "topicalization" will be discussed in a later section. Myhill also remarks (1983:161) that the position of "marked" focus is "preverbal", as opposed to "unmarked" focus which is "postverbal": "focusing places arguments in preverbal position and is a marked construction".

That there exists in Tzotzil a difference between "marked" and "unmarked" focus may be confirmed by consideration of wh-questions and their answers. Unfortunately, there are very few examples, in the data I have, in which the entire sentence is repeated in an answer. All such question and answer examples I have are more like the form in (113):
113) a. k'usiy- o t- al ch- a- pak' an- be
what A3 tortilla poss icp A2 make io
'What are you making tortillas for?'

b. yotal ve7elil
'tortillas for dinner'

114) a. bu ch- a- bat lol
where inc 2abs go Larry
'Where are you going, Larry?'

b. ch- i- bat ta jobel
inc labs go to San Cristobal
'I'm going to San Cristobal'

(Aissen 1987:164-165)

The one example in which the entire sentence is repeated in the answer is as follows:

Myhill (1983:161) presents this example to indicate that "answers to questions are not normally focused" in Tzotzil. He writes that "The answer to a question is focused; this choice is usually neutral and hence receives unmarked focus marking but not marked focus marking".

Myhill and an awareness of discourse studies on other Mayan languages which precludes the possibility in the present work that Tzotzil might be consistently RHEME-initial, like Hixkaryana. If: 1) the placement of NEW information in response to wh- questions is a good diagnostic of RHEME's prototypical placement in a language (as was claimed in Chapter Four); 2) such NEW information is not sentence-initial in Tzotzil; and 3) sentence-initial placement is utilized in Tzotzil for discourse emphasis; then, some factor other than "RHEME-initial" must avail itself for an explication of Tzotzil word order. The explanation beign offered here is that Tzotzil is consistently DIFFUSE-initial.
This is to say that a cognitively "active" or "exciting" semantics is consistently associated with sentence-initial position in Tzotzil. The discussion of Tzotzil thus far has presented claims by researchers that the language is "comment-topic" in basic order; it has presented constructions in which elements are "fronted for emphasis", and introduced the emphatic particle *ja7*, noting its sentence-initial placement; it has discussed variations in word order with quantitative expressions, in which S's and O's and/or their quantifiers are fronted for emphasis of some kind; and finally, it has presented Myhill's suggestions that both "marked" focus and topic occur sentence-initially, while their "unmarked" analogs do not. All these observations are consonant with a cognitively active, DIFFUSE semantics accruing sentence-initial position in the language.

5.5.2. Definiteness and Sentence-final Position

It has been noted by researchers on Tzotzil that sentence-final position is normally associated with elements which are definite. Aissen (1983:275) says that Tzotzil *ti* and *li* "function somewhat like definite articles and differ in proximateness, *ti* being the less proximate". Haviland (1981:30-31), more specifically, calls both *ti* and *li* definite articles, saying that "*li* is the definite article which signals the proximity of a specific thing", while *ti* indicates that the noun is "definite and specific" and that its referent is "remote or distanced in time or space". Both Haviland (1981:30) and Aissen (1983:275) note that a definite noun is preceded by one of these articles, and followed by the enclitic -e. Indefinite reference is expressed by the noun alone, with no article (Haviland 1981:30). For example:

115) a. vo7  
     'water'
Aissen (1983:275) notes that the enclitic -e occurs "in phrase-final position".

Thus there are morphological indications of the degree of definiteness or specificity of a nominal; what is important to note is that position within the sentence also comes into play in this regard, as explicitly discussed by Haviland (1981:33-39). Consider the differing positions of vo7 'water' in (116), and 7ixim 'maize' in (117):

116) 7oy vo7 ta k'ib
    there-is water def-prep pitcher
    'Hay agua en el cantaro'
    'There is water in the pitcher'
    (Haviland 1981:33)

117) 1i7 7oy li 7ixim- e
    here there-is the maize cl
    'Aquí está el maíz'
    'The maize is right here'
    (Haviland 1981:33)

In (116), vo7 'water' is nonspecific and indefinite, but its location is specified: the order of nominals is "indefinite-definite". In (117), 7ixim 'maize' is specific and definite and as such occurs in final position. The location in this sentence is also definite (1i7 'here'); Haviland says that 1i7 has been "fronted" (1981:33) in this sentence. Thus, even though 1i7 indicates a specific place, and thus normally would occur near the end of a sentence, the location is being emphasized and thus occurs sentence-initially; this emphasis is reflected in both the Spanish and English glosses.

Haviland has noted that "The most definite noun in each sentence occupies
final position" (1981:37). He gives the following examples:

118) (te) 7oy 7ixim ta moch
demonstr. there-is maize def-prep basket
'(Ahi) hay maiz en la canasta'
'(Over there) there is maize in the basket'

119) te (7oy) ta moch li 7ixim- e
demonstr. there-is def-prep basket the maize cl
'El maiz esta en la canasta'
'Over there ([there is]) the maize in the basket'

(Haviland 1981:37)

He comments concerning these sentences as follows:

In the sentence indicating existence [118], the noun 7ixim 'maize' is indefinite -- it means 'some maize, an unspecified quantity of maize'; 7ixim takes no articles, and remains in middle position. On the other hand, .... ta moch 'in the basket', being the most definite constituent of the sentence, goes in final position .... [In (119)], the subject -- the noun in final position -- li 7ixime is very definite: 'the maize' (for example, 'the maize we are going to eat').
(Haviland 1981:37)

The compatibility of definiteness with sentence-final position is also evident from the sentences in (120):

120) a. 7i- s- pet 1ok'el 7antz ti t'ul- e
cp A3 carry away woman the rabbit cl
'The rabbit carried away the woman'

b. ? 7ispet 1ok'el ti 7antz ti t'ule
'The rabbit carried away the woman'

(Aissen 1987:1, 3 and Aissen 1983:275)

Aissen (1987:3) comments that (120a) is "grammatical", while (120b) is "awkward". In (120a), t'ul 'rabbit' is the Subject, and definite in reference; as such, it occurs with the definite article ti and the phrase-final clitic -e.

Significantly, it also occurs in sentence-final position. The Object in (120b), 7antz 'woman', bears a definite gloss in the English, but the Tzotzil does not carry a morphology of definiteness; and the addition of the definite article ti to
the Object is felt to be "awkward". Haviland (1981:55) notes specifically that

Zinotec Tzotzil avoids sequences of nouns with the
definite article. In a sequence of definite nouns, only the
last noun takes the article.

He gives the following example:

121) a. te ta sna Xun li 7ixim e
   there def-prep house Juan the maize cl
   'El maiz esta en la casa de Juan'
   'The maize is in Juan's house'

b. * te ta sna li Xun e li 7ixim e
   (Haviland 1981:55)

The strong attraction of the expression of definiteness to sentence-final
position is evident.

Thus, according to Haviland's interpretation, and to Aissen's statement of
the occurrence of 7a liiti ... -e (the definiteness "frame"), there is a strong
tendency in Tzotzil for definiteness to be associated with sentence- or phrase-
final position. This is significant for understanding why the language is OS in
its basic word order, since Subjects are more likely than Objects to be definite
in reference (see Givon 1984:422). Haviland's statement that Tzotzil is
"Comment-Topic" in functional order is also consonant with these
observations. The correlation of Subject and Topic is a common, though not
universal tendency cross-linguistically (see Chapter Six). Grammatical
definiteness, syntactic Subjects, and functional TOPICS are all compatible with
the cognitively "stable", "unexciting" semantics characteristic of psychological
FOCUSEDness. Expressions of this psychological FOCUSEDness in Tzotzil
consistently occur in sentence-final position.

5.6. Alternate Word Orders
5.6.1. Introduction

Though the basic word order in Tzotzil is considered to be VOS, the following orders also occur: SVO, OVS, OV, VO, SV, VS. In addition, there is a construction involving left-dislocation, which results in the following orders: S,VO; S,V; O,V. The function of these alternate word orders will be discussed in this section.

It has already been noted that sentence-final position in Tzotzil is reserved for the expression of "unmarked topics" and of S's in basic word order. The attraction of the expression of definiteness to sentence-final position has also been discussed. Examples of SVO order were presented earlier ([108] through [112]), and note was made of the RHEME-like character of the English glosses. The emphatic sentence-initial placement of modifying phrases, which would normally occur near the end of the sentence, has also been discussed.

5.6.2. Left-dislocation

S's, O's, and modifying phrases may all be left-dislocated in Tzotzil. Elements which are left-dislocated occur in the "frame" 7a li/iti ... -e. Aissen (1987) refers to the particle 7a as a topic-marker, and glosses it as such. Haviland (1981:55) simply refers to it as an "introductory particle". I will present a suggestion as to the function of the sentence-initial particle 7a below.

Both Aissen and Haviland note the necessity of definiteness in the sentence-initial construction involving the "frame" 7a li/iti ... -e. Aissen writes: "In general, a nominal dependent of the verb may occur in preverbal position if it is definite. Topicalized nominals frequently occur in the frame: 7a li -- e"
Haviland comments that the "introductory particle 7a goes at the beginning of sentences with fronted definite nominals" (1981:55). He also notes that "the subject of a sentence may be fronted -- being moved to the beginning of the sentence in order to give [it] greater emphasis" (1981:54).

Examples of S's occurring in left-dislocated constructions (with the introductory particle 7a) are provided by sentences (122) through (125).

122) 7a ti tzeb- e, 7i- s- sa7 s- mala1 topic the girl cl, cp A3 search A3 husband 'The girl looked for a husband'
      (Aissen 1987:18, from Laughlin 1977:80)

123) 7a li Xun- e, l- i- y- ak'- be s- kalto- al topic the Xun cl, cp B1 A3 give io A3 broth poss 'Xun gave me its broth'
      (Aissen 1987:138)

124) 7a li chon- e, mu s- k'an mil- el topic the snake cl, not A3 want kill el 'The snake doesn't want to be killed'
      (Aissen 1987:16, from Laughlin 1977:166)

125) 7a ti rey- e, 7i- 7ak'- b- at topic the king cl, cp let io psv
    la s- t'uj jun y- aqnil cl A3 choose one A3 wife
    'The king was allowed to choose a [one, EAC] wife'
    (Aissen 1987:218)

It is possible, in some cases in the data I have, to determine the discourse function of S's occurring in this construction. Sentence (126), for example, seems to involve change of TOPIC:

126) There was a man and a woman, newlyweds, --

    7a ti vinik- e, ta x- lok'... topic the man cl, icp leave...

    7a ti 7antz- e, jun yo7on ta x-kom topic the woman cl, one heart icp remain
    'The man, he goes out ... the woman, she stays at home happily'
That is, it may be that in cases like (126), (127b) and (128b), what is involved is re-instatement of a KNOWN PARTICULAR as TOPIC:

127) a. te ta Jobel li Xunka7 e there def-prep San Cristobal the Juana cl 'Juana esta en San Cristobal' 'Juana is in San Cristobal'

(Haviland 1981:54)

b. 7ali Xunka7 e , te ta Jobel 7a-the Juana cl, there def-prep San Cristobal 'Y Juana, ella esta en San Cristobal' 'And Juana, she is in San Cristobal'

(Haviland 1981:54)

128) a. 7oy smalal li Xunka7 e there-is husband the Juana cl 'Juana tiene esposo' 'Juana has a husband'

(Haviland 1981:54)

b. 7ali Xunka7 e , 7oy smalal 7a-the Juana cl, there-is husband 'Hablando de Juana, ella tiene esposo' 'Speaking of Juana, she has a husband'

(Haviland 1981:54)

Sentence (129) provides an example of a temporal occurring in the left-dislocated construction:

129) 7a ti vo7ne- e , 7o la jun vinik ... topic the former times cl, there-is cl a man ...
    'Long ago there was a man ...'

(Aissen 1987:18, from Laughlin 1977:52)

It has been mentioned that O's may also be "fronted", and left-dislocated in this manner. Sentences (130) and (131) provide examples of this:

130) 7a ti s- tot ti tzeb- e , topic the A3 father the girl cl,

7i- k- il- be ta jobel

cp Al see io in San Cristobal
'I saw the girl's father in San Cristobal'

(Aissen 1987:155)

131) 7a  l- a- tot  li  vo7ot- e ,  l- a-  j- nup- be  ta  be topic the A2 father the you cl, cp B2 A1 meet io on road 'I met your father on the road'

(Aissen 1987:155)

The next example (132) gives more possibility of suggesting the discourse factors affecting the occurrence of sentence-initial O's, since it has more context given:

132) Something had landed at the foot of the tree, they went to look. There was a straw mat. "Hell, what could it be? Come on, let's untie the straw mat!" the two men said to each other. They untied it. You know what? --

a. tzeb san- antrex 1a te s- ta- ik 7un
girl San Andres cl there A3 find pl cl
'They found a San Andres girl there'

b. 7a  ti  tzeb san- antrex 7une , 7i- y- ik'- ik 1a ech'el 7un
topic the girl San Andres cls , cp A3 take pl cl away cl
'They took the San Andres girl with them'

They lit a fire for her, they let her warm up, they gave her tortillas, they fed her, they slept with her.

(Aissen 1987:158-159)

In (132a) the word order is O-locative-V. This sentence is interesting for several reasons. First, note that the O is not left-dislocated, but simply sentence-initial. The O tzeb san-antrex 'San Andres girl' carries an indefinite gloss in the English, consonant with the observation which has been made that indefiniteness is incompatible with sentence-final position. Second, note that the O constitutes NEW information, in response to a question. Although the question is not a "wh- question", still the O may be construed from the context to be of this NEW discourse status, and hence is an example of "marked" RHEME. In sentence (132b), the O has been left-dislocated and occurs in the
definiteness "frame" 7a ti ... -e. In this sentence, the O constitutes a change of TOPIC.

It is becoming clearer that O's are not consistently associated with semantic REMOTEness in Tzotzil, as they were in Hixkaryana. Recall that TOPICs are characterized by a semantic IMMEDIATEness and are psychologically FOCUSED, while elements bearing RHEME are characterized by a semantic REMOTEness and are psychologically DIFFUSE. The sentences in (132) are important in this regard because they provide examples of O's apparently being associated first, as NEW information, with a semantic REMOTEness (132a), and then, as TOPIC, with a semantic IMMEDIATEness (132b). Thus, in Tzotzil, unlike in Hixkaryana, O's may function as TOPICs. The further significance of the sentences in (132) will become apparent just below, in the discussion of the introductory particle 7a.

5.6.3. Definiteness, Indefiniteness, and the Particle 7a

Recall that there is a strong tendency in Tzotzil for nominals which are definite to occur sentence-finally. In the (sentence-initial) left-dislocated constructions just presented, however, nominals which are left-dislocated necessarily occur within what I have called a "frame" of definiteness: 7a li/ti ... -e. Recall also that Aissen (1983, 1984a, 1984b, 1987) has considered 7a to be a marker of "topic", and has glossed it as such in her data. A different, more comprehensive suggestion for the function of 7a will be offered here.

Definite nominals either occur in their "normal", postverbal positions, or they occur in a sentence-initial, left-dislocated position in a phrase introduced by 7a. Recall that in basic order, both S and O would be postverbal. Haviland
(1981:56) notes that "indefinite nouns also can be fronted, but do not require the initial particle 7a". He notes that the fronting of indefinite nominals is "somewhat rare" (1981:56). He gives the following examples of indefinite and left-dislocated nominals:

133) a. ("normal order")

7oy 7ixim ta kuveta
there-is maize def-prep bucket
'Hay maíz en la cuberta'
'There is maize in the bucket'

b. te ta kuveta, 7oy 7ixim
there def-prep bucket, there-is maize
'En la cubeta hay maíz'
'In the bucket there is maize'

c. 7ixim e, 7oy ta kuveta
maize cl, there-is def-prep bucket
'Maíz hay en la cuberta'
'Maize is in the bucket'

(Haviland 1981:56)

Sentence (132b) is also an example of a "fronted" indefinite nominal; though not left-dislocated, it is not introduced by 7a.

The present interpretation of sentential word order in Tzotzil has suggested that the language is DIFFUSE-initial and FOCUSED-final. If this is accepted, then the expectation at the beginning of each Tzotzil sentence (or phrase) would be for a discourse element which is semantically RHEME-like, cognitively "active", new in some sense in the onward movement of discourse. If the sentence-initial element does not confirm this expectation of psychological DIFFUSness, but rather is psychologically FOCUSED, this is contrary to discourse expectations, and requires a formal mark in the language. This formal warning of (unexpected) discourse continuity rather than (expected) discourse discontinuity is carried out by the introductory particle 7a. Syntactic marking having a similar psychological function was
discussed in relation to RHEME-initial Hixkaryana.

5.7. Concluding Remarks on Tzotzil

While evidence has been presented that Tzotzil is DIFFUSE-initial and FOCUSED-final, O's in the language are not consistently tied either to the semantics of REMOTEness or to that of IMMEDIATEss, unlike the O's in Hixkaryana. In Tzotzil sentences in which both O and S are nominally expressed, the S is expressed finally because it is the most definite constituent. This is compatible with a psychological order of DIFFUSE-FOCUSED. Likewise, the sentential verb in such sentences is the element most compatible with the psychological DIFFUSEness associated with sentence-initial position in the language. The O in such sentences falls rather "in-between" on the scale of DIFFUSE-FOCUSED, and this is mirrored in the syntax of these sentences.

Thus, like other Mayan languages, definiteness in Tzotzil affects word order. This tendency in Mayan is noted by Quizar (1979:84) in her discussion of the problems associated with determining basic word order in the family:

Features such as relative animacy, definiteness, and specificity of the two noun phrases, have all been found to affect word order in Mayan languages. Not all of the languages that have several morphologically unmarked word orders have been studied in great detail, so that some of the ways in which these and other factors affect word order in Mayan are still unknown.

Brody (1984a:65ff.; 1984b:720ff.) and Pinkerton (1978) provide detailed discussions of this matter with regards to Tojolabal and K'ekchi, respectively.

5.8. "Basic Word Order", Hixkaryana, and Tzotzil
The interpretations of Hixkaryana and Tzotzil were initially begun by a concern the syntactic ordering of S's, O's and V's. During the course of investigation, and the subsequent telling of the stories, these syntactic constituents did not manifest the primal importance that the Greenbergian typology seems to give them. Whether an element was a Subject, Object, or Verb, or at times, whether it was a Noun or a Verb, did not appear as relevant to an understanding of sentential order as did functional, pragmatic, and psychological considerations.

Hixkaryana and Tzotzil constitute a serious challenge to the concept of "basic word order", and to the typological classification which this concept has enabled. It appears that other Amazonian and Mexican languages present similar challenges. For discussions of Amazonian languages and their implications, I refer the reader to the following publications: Derbyshire and Pullum 1986b; Harrison 1986; Payne 1986; Howell 1983; and Hoff 1978. For discussions of Mayan and Mexican languages, the following publications are relevant: Brody 1984b; Pickett 1983; Pinkerton 1976b, 1978; Walter 1980; and Quizar 1979.

Before moving into a discussion of the broader significance of the findings of my language interpretations both for linguistic typology, and for linguistics in general, there is one salient point concerning Mayan languages which must be mentioned.

It was noted in the discussion of Hixkaryana that the proper kind of sentences for determining basic word order were actually quite rare in discourse: simple transitive sentences with both the S and the O nominally
expressed. It appears that a similar situation occurs in Mayan, as noted by Quizar (1979:71):

When going through texts of various Mayan languages to look at word order, I had a great deal of difficulty finding transitive sentences.

Brody (1984b:719) notes that in Tojolabal discourse, "sentences with one argument ... are actually the type of sentence which occurs with greatest frequency ... where context disambiguates". She continues:

Thus, by most standards, VOS is the basic word order of Tojolabal. However, VOS sentences are difficult for hearers to process when they occur in isolation, removed from any context which would assist in interpretation .... Calling VOS sentences 'difficult to process' does not mean that these sentences are marginally grammatical; they are good, fully grammatical sentences. The difficulty lies in making the correct interpretation. Both NPs are potential candidates for either the grammatical role of subject or that of object, since they are both third person.

In a similar vein, Pickett (1983:531) notes that in the sixty-two Mexican Indian languages she surveys,

Many languages have the verb, with its bound subject and/or object, as the only obligatory element of a clause. Nouns, as optional elements, are used only when necessary for clarity of information, and free pronouns are used only for special emphasis. In some it is rare to include both nominal subject and object in the same clause ...

Finally, Laughlin (1977:6), writing of the difficulties in translating his Tzotzil texts into English, says:

To compound the translator's problems, subjects and objects may become as invisible as the Cheshire cat. Thus in the Tzotzil word smaj we know that the subject is he, she, or it, the root is hit, and the object is he, she, or it. The precise identity of the beater and the beaten is understood ... Furthermore, if the plural marker -ik is added to produce smajik not only do we remain in the dark as to whom the subject and the object are, but we do not even know whether there are many aggressors and one victim, or many aggressors and many victims. Worse yet, once it is understood by the Tzotzil exactly who is in the
plural they feel no compunction about dropping off the plural marker so that the translator discovers that a crowd has shrunk to one with no apparent explanation!

These concluding remarks will serve as an entrance into the next chapter.
CHAPTER SIX

TAKE IT OR LEAVE IT

6.1. A Puzzle

It has largely been accepted that the character of human apprehension is such that what is stable and able to be grasped because known is most important to, and most attractive to, focal awareness. It is a common teaching technique in the West to start a pupil learning new things by beginning with what he knows already, and where the landmarks are familiar, and then by proceeding to newer territory. This teaching technique may be contrasted with one associated more with non-Western cultures, in which a pupil is presented with a puzzle, or a paradox, which is to be contemplated. The learning experience is not so much getting a fact in (the) mind, like a thing into a container, as a mulling over, and acceptance of, the simplicity of a maze. Rather than the new being grasped and finding its place in the comfort and stability of what is old, the old is contemplated.

The Indo-European grammatical tradition, a heritage or cultural artifact principally of Indo-European speakers, has been one which has assigned labels to bits and pieces of languages which have been made visible by literacy. These labels have included "parts of speech", like Nouns and Verbs, and "parts of the sentence", like Subjects and Objects. The syntactic typological movement instigated by Greenberg (1963) has operated within this inherited grammatical domain of analysis. The labels of Subject, Object and Verb have enabled this categorization of the languages of the world.
6.2. How Can a Language Be "Rare"? On the Concept of Markedness

The statistical rarity of OS languages is certainly striking. To say that something is "rare" does not necessarily carry the connotation of "oddity" and "strangeness" which often appears in the discussion of these languages. The oddity is sometimes cast in terms of "poorly attested", or "poorly understood": that is, once these odd languages have been understood properly, they will not seem so odd. In other words, the languages some day will (be made to) fit into existing theory. Another connotation of "rare", however, is that of something's "being precious, because scarce"; accompanying this connotation is a feeling of exhilaration, like the breathing of (rare) mountain air. What then has enabled the connotation of "rare-as-oddness" to be associated with OS languages? How can a language be classified as "rare"?

The typological discussion of languages in this century has taken the concept of "markedness" as a methodological assumption. An understanding of the questions just posed will be eased by a discussion of this concept and its practical implementation in the determination of syntactic typology.

The manner in which linguistic data is interpreted in order to ascertain "basic word order" is fully dependent on a notion proposed by Trubetzkoy of the Prague school in his work in phonology: the notion of "markedness". This notion has provided an organizing principle for the neat classification of phonemes and, in word order typology, of syntax. This principle has been employed by linguists in this century, perhaps sometimes only as a "heuristic device", but most often as an organizing principle for the construction of theories and diagrams of languages: that is, for their visual classification. It
is beyond the scope of this discussion to fully explore this curious principle and the history of its development, but its deployment in the determination of "basic word order" must be touched on.

The concept of "basic", "dominant", "unmarked", or "neutral" implies, and indeed strives to establish, a norm against which may be gauged occurrences of what is "nonbasic", "non-dominant", "marked", and "non-neutral". This norm of what is "basic" serves several purposes.

It allows for the simultaneity and flow of speech to be broken down into stable, individual constituents, which enables the (literate) narrative of the classification of languages to begin. It is an arbitrary [forced choice] decision as to which "constituent order" is claimed as "basic", and then the story can begin (see Pickett 1983; Brody 1984b). This classificatory story is not one that takes place in the everyday, practical world (cf. Tyler 1978; Luria 1976), but is one that takes place within an attempted scientific discourse, hampered by the exigencies of the technologies of literacy. Once posited, the "basic", "dominant", "neutral", "unmarked" form serves both as the standard against which other forms are measured, and as the source from which these other forms emanate. Here, once again, the old metaphors of measurement and Platonic emanation can be felt to be working. The basic form as "standard" is considered to be more fundamental than other forms, which are "variations", or "derivations". The basic form as "source" is considered to be somehow more healthy, pure, and normal than other forms, which are "exceptions", "deviations", or "marked" (cf. "scribbled on") constructions. Another old metaphor which is playing in the terminology for this concept is that which stems from the ideal of scientific discourse, an ideal which seeks to remove any
voice of person from its wordings. Upon what grounds can it be claimed, at any level, that what is "neutral" (cf. "of indeterminate color", "out of gear", "taking no sides") is fundamental or "basic" to anything that has to do with people?! Natural scientists had plenty of trouble finding any thing in the physical world which could be characterized as "neutral" or "basic" or stable; the fact that in linguistics (a discipline of the "human sciences") such a notion is still alive seems just a trifle quaint.

As well as these considerations concerning the very concept of "markedness", its practical implementation as a diagnostic or classificatory tool is often quite thorny. There are what might be called "methodological" problems with the determination of basic word order, and they are serious. These involve the questions of what is appropriate data, the manner in which that data is acquired, and the manner in which that data is interpreted. Each of these considerations are complex, and merely pointed out here.

In providing what they believe constitutes appropriate data, linguists have long been concerned with the intricacies of phonology, morphology, and syntax. Much valuable information has been garnered, but it has led to the postulation of descriptive and classificatory theories for these various levels and (sometimes) their inter-relations. Linguists have largely been stuck inside sentences. The grammars and papers they write, which provide knowledge of far-flung languages to the research community, reflect this limitation. In addition, the postulation and perfection of theoretical frameworks has come to dominate linguistic discussion; in reading such discussions one gets more a feeling for the theory than for the language or languages they purport to (re-)present. (In addition to this, there is the Indo-European grammatical
basis for linguistic theories which has been alluded to frequently in the present work.)

What is the nature of language "data" beyond the sentence? There are at least two kinds which appear in the literature: stories and conversations. Both of these are tape-recorded and then written down, so that their content may hold still. This recording and transcribing of stories and conversations between informants yields "texts" which then may be posed the questions the linguist asks. This is another complex area, however. It touches on the relationship of orality and literacy, for instance. If heard conversations and heard stories are thus transformed and translated into texts, how can one maintain the position that one is a student of "speech"?

Thus the "discourse" linguist most often works with texts. If one's task is to seek and find in a foreign language a certain sentential word order which one can report as being "basic" to that language, the task is complicated by the question of genre. Story-telling differs from conversation-making, both of which differ from story-writing and dialogue-writing. To either claim that one of these language situations is prime, or to pick sentences from them indiscriminately for the sake of one's theoretical analysis indicates more the power of the urge to classify and theorize than it does sensitivity to language situations or to a specific language.

Textualized stories and conversations are two genres, then, from which linguistic data can be taken for the purposes of word order investigation. There is a third situation, that traditionally made use of by field linguists, which I suggest can be also considered a genre: elicitation.
The genre of elicitation is created in the process of acquiring linguistic data. If one's interest is to determine a "basic", "neutral" word order for a language (in conformance with one's theoretical -- or perhaps pre-theoretical -- expectations), one can attempt this through the practice of elicitation. Now, the questions one asks will often determine the answers one receives; in the context of the elicitation setting, linguists are particularly susceptible to suffering from this fact. (For discussion of this with specific reference to word order investigation, see Pickett 1983:532-533; and Brody 1984b:719.) In addition, the elicitation situation is a peculiar one to be engaged in at all (for both parties) because there are uncomfortable interpersonal dynamics: employer versus employee; socially dominant versus socially inferior; highly educated versus (often) poorly educated; literate versus (often) illiterate/nonliterate; knowledge-maker versus data-giver. This list could continue. These peculiar social dynamics are the background and source of the (textualized) elicitation genre. A third peculiarity of this research method is that elicitation is usually bilingual, and carried on in a socio-linguistically dominant language (the Indo-European language of the linguist). For the elicitation of basic word order, this causes at least two problems.

The first is the influence (the "interference") of the language of questioning upon the language of answering. Pickett (1983:531) notes this kind of interference, with specific reference to syntactic word order:

In Mexico, the structure of Spanish often influences the structure of elicited indigenous language material. My students found that using elicitation in Spanish, an SVO language, often resulted in higher frequency of SVO order in the data elicited. Bilingual indigenous translators often show this same influence in translating Spanish materials into their own languages, resulting in more SVO sentences than normal.
Brody (1984b:719) notes the problem of bilingual elicitation in word order study by referring to differing pragmatic and semantic meanings in the two involved languages:

work on Tojolabal and any of the Mayan languages is usually done through Spanish, a language which is also characterized by a variety of word orders, each of which carry different semantic and pragmatic meanings, and there is no reason to suspect that, for example, SVO in Spanish is precisely equivalent to SVO in Tojolabal.

Secondly, there is the frequent necessity of translating the glosses, from the language of questioning into the language of publishing, often English. For example, Carib and Mayan elicitation is done through Portuguese and Spanish. In order to be published, and to reach a wider audience, translation of the Portuguese and Spanish glosses into English is required. While Portuguese, Spanish and English are all Indo-European languages, there are obviously large differences in the nuances of meaning conveyed by word order in these different languages (as just noted by Brody) -- nuances which are lost in translation.

Thus, the elicitation situation is one which is peculiar unto itself; the texts which it generates can be considered a genre which is created in the acquiring of linguistic data.

The concept of "markedness", and the determination of "basic word order" for a specific language which this concept enables are thus both problematic. The quest for "linguistic universals" has, through the medium of word order typology and its Indo-European grammatical basis, in effect posited a universal word order in which the Subject precedes the Object: certainly the statistics tend to this conclusion. Thus languages in which the Object precedes the
Subject in "basic" order are not only "rare" statistically, but universally "non-basic", "marked", and "exceptional" word orders.

6.3. On the "Unnaturalness" of OS Languages

Because of the preponderance of SO word orders, it has been assumed that this order is somehow more "basic", "fundamental", "natural", or "human". However, there is a striking difference which has been shown to operate, at the sentential level, between languages which are considered to be SO and those considered to be OS.

To accentuate this difference, it is necessary to call to mind (in a very cursory fashion) attempts to characterize the notion of 'Subject'. Halliday (1985:32-37) discusses 'Subject' on a functional basis, and proposes the terms "Theme", "Subject", and "Actor", respectively, for what, he says, came to be known in the second half of the nineteenth century as "Psychological Subject" ('that which is the concern of the message'), "Grammatical Subject" ('that of which something is predicated'), and "Logical Subject" ('doer of the action') (see also Firbas 1974:9, 11). Halliday notes that in some languages these functions are kept distinct, and in others they are "conflated". Here I will be concerned with what Halliday calls the function of "Theme" (the "Psychological Subject"). Halliday writes that

The Theme is the element which serves as the point of departure of the message; it is that with which the clause is concerned. (1985:38)

After "one element in the clause is enunciated as the theme", it "combines with the remainder so that the two parts together constitute a message"; the "remainder of the message, the part in which the Theme is developed, is called
in Prague school terminology the Rheme" (Halliday 1985:38). Halliday notes
the usage of the terms "Topic" and "Comment" instead of Theme and Rheme,
but claims that "the label 'Topic' usually refers to only one particular kind of
Theme" (1985:39). (The terminology of TOPIC and RHEME used in the
present study has been that of Davis [MSS. a, b, c], and is somewhat
analogous to these terms; the specific characterization of TOPIC and RHEME
employed by Davis was discussed in Chapter Three.)

What is important for the present discussion is not a resolution of
terminology, but to point out that the analogous notions of Psychological
Subject/Theme/Topic have been noted by linguists to correlate with a certain
sentential or clausal position. Halliday does not make claims for universality;
however, he does state

As a general guide, the Theme can be identified as that
element which comes in first position in the clause. We
have already indicated that this is not how the category of
Theme is *defined*. The definition is functional[.].... it is
what the clause is going to be about. (1985:39)

He comments that this is somehow a "natural" organization of a "message":

if in any given language the message is organized as a
Theme-Rheme structure, and if this structure is expressed
by the sequence in which the elements occur in the clause,
then it seems natural that the position for the Theme
should be at the beginning, rather than at the end or at
some other specific point. (1985:39)

The Rheme, or "remainder", may come in clause-initial position, but Halliday
calls this a "marked" construction:

in a Theme-Rheme structure it is the Theme that is the
prominent element contrast with the typical pattern there
is a standing-out or MARKED alternative, ... in which the
usual relationship is reversed. (1985:42)
Hoff (1978) directly aligns himself with the Prague school "functional sentence perspective", and discusses these terms in an analogous manner:

> [the] concepts theme and rheme to some extent are counterparts of topic and comment, but differ from these mainly by their relative and gradual character. Theme and rheme mark the ends of a continuous scale: in an actually spoken utterance, meaningful elements that make a relatively smaller contribution to the process of communication are nearer to the thematic end of this scale, while elements that make a relatively larger contribution are nearer to the rhematic end. (1978:16)

Hoff (referencing Firbas [1974:22] and Vennemann [1975:288]) re-iterates the relative position of theme-like and rheme-like elements:

> Unless some different order is imposed on them by syntactic rule, meaningful elements tend to assume a linear sequence in which elements nearer to the thematic end of the scale precede elements nearer to the rhematic end of the scale. (1978:17)

He also notes the possibility of the "marked" order of Rheme-Theme; following Prague school linguists, he calls this "the 'pathetic' or 'emotional' order rheme-theme which contrasts with the 'natural', 'neutral', or 'stylistically unmarked' order theme-rheme" (Hoff 1978:17).

Linguists other than those of the functionalist persuasion have noted this same "theme-rheme" construction of clauses and have assigned this same order as the "natural" order. Hockett (1958:201) does not explicitly state that theme-rheme or topic-comment is the "natural" order of these elements, but he too notes the strong tendency for this order:

> The kernel of an English sentence of the favorite sentence-type is a predicative constitute. This is true also in most other languages, and quite possibly in all, though there are subsidiary differences .... The most general characterization of predicative constructions is suggested by the terms "topic" and "comment" for their ICs [immediate constituents]: the speaker announces a topic and then [my italics, EAC] says something about it.
Greenberg (1963:61), of course, also notes this quasi-universal tendency in languages, though he phrases it in terms of Subject and Object:

Universal 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.

Chomsky (1965:221) writes

It might be suggested that Topic-Comment is the basic grammatical relation of surface structure corresponding (roughly) to the fundamental Subject-Predicate relation of deep structure. Thus we might define the Topic-of the Sentence as the leftmost NP [sic] immediately dominated by S in the surface structure, and the Comment-of the Sentence as the rest of the string. Often, of course, Topic and Subject will coincide.

(Note the terminology of "leftmost NP" in this quotation from Chomsky, and in the one from Keenan, below: examples of the paperification of language.) More recently, Keenan (1976b) has made an attempt towards a "universal definition of 'subject'". He proposes that "subject" is a "multi-factor concept", and lists more than thirty properties which subjects might possess (the trouble Keenan has indicates the questionability of the universal definition he seeks). One of his subject properties reads as follows: "b[asic]-subjects are normally the leftmost occurring NP in b[asic]-sentences". He specifically notes, however, that other orders do exist:

in a few cases languages have fairly fixed word order in which the subject follows one or more objects .... [and] languages ... in which basic word order appears totally free, and languages ... in which NPs in b[asic]-sentences occur in any order as long as they are all after the verb. (Keenan 1976b:320-321)

The listing of similar statements could go on. It is simply a commonly held assumption that the Psychological Subject/Theme/Topic occurs in sentence- (or clause-) initial position, and that this order is somehow "natural" for human
understanding. If a sentence is Object-Subject, rheme-theme, Comment-Topic in its word order, it is considered to be a "marked" construction; this eliminates it as a candidate for the "basic" or "natural" order in a language.

On the other hand, it has been noted by researchers in OS languages that this supposedly "natural" and perhaps universal claim for Theme-Rheme in basic word order does not hold for these languages. What is to be made of this?

6.4. A Brief Deconstructive Moment

The possibility of Greenberg's (1963) seminal paper relies on at least four distinct assumptions. What Greenberg proposed as tendencies and suggestions (while, paradoxically, calling them "universals") launched a scrambling by researchers to make his suggestions and observations into a unified theory of linguistic universals. This scrambling, of course, involves the initial assumption that "unified theory" is a possibility, and that the construction of such a thing is a useful, interesting, or at least valid, enterprise.

A first assumption is that grammatical Subjects, Objects and Verbs have universal validity in the world's languages. This assumption is effectively demonstrated to be untenable in several recent publications (Schachter 1976; Li and Thompson 1976; Foley and Van Valin 1977; Pickett 1983; Dik 1983; Brody 1984b; Hopper and Thompson 1985; Harrison 1986).

A second assumption is that a linguist can objectively and clearly determine one sentential word order in a language as being "basic". The untenability of this assumption has also received attention in the literature (Pickett 1983; Brody 1984b; Payne 1986), and has been touched on in this
chapter, in the discussion on "markedness". Its untenability involves serious methodological and theoretical problems, as well as the uncomfortable fact that language data simply do not behave properly.

A third assumption in syntactic typology is that the unit of the sentence is the long-sought-for 'type' (see Chapter Two) upon which inter-linguistic comparison and classification may fruitfully be conducted. The untenability of this third assumption on a universal basis stems in part from the untenability of the first assumption discussed. That is, all languages must "have" Subjects, Objects and Verbs, and that only these are relevant constituents to expect in a language. In Hixkaryana, and in the Mayan and Amazonian OS languages I have looked at, it is simply misleading to seek the data needed for the ascertaining of basic word order: transitive sentences with nominally expressed Subjects and Objects. In these languages, such sentences are very rare in narratives and discourse, and to elicit these sentences for the purpose of "telling the folks back home" is to misrepresent the nature of the languages. To use the terminology of syntactic typology, transitive sentences of this type in these languages are "marked constructions"; thus to be forced to use this as a basis upon which to compare them with other languages is quite ironic.

It is perhaps appropriate at this point to simply mention that is is very curious that, according to syntactic typologists, transitive simple sentences are de facto more "basic" than other sentence types (see Brody 1984b:713).

The assumption that the sentence is an appropriate 'type' for universal cross-linguistic comparison is also belied by the nature of Hixkaryana discourse. What Derbyshire calls "sentence clusters" in the language involve "a
string of two or more sentences with near identity of semantic content" (1977b:176). Hixkaryana discourse, both in stories and in conversations, is apparently not structured by units of meaning composed of sentences. (It is beyond the scope of the present work to further pursue this; but see Derbyshire [1977b] for an initial discussion.)

A fourth assumption in syntactic typology is that the construction of a unified theory of universals based on syntactic word order is a valid enterprise. This assumption is simply shattered by the untenability of the first three assumptions.

The quest for some basis upon which languages can be universally compared, then, has still not been answered. No satisfactory ‘type’ has been found tenable which will allow for a neat classification of the world’s languages. This leads one to question the question. Is there a possibility of typological classification on any basis? Should the valuable energies of researchers continue to be directed towards the quest for ‘type’, and towards the construction of a classification into which all languages may be stuck, or neatly placed?

It was intimated earlier that the notions of ‘type’, and of ‘theoretical classification’ are both constructs of the analyst. Hence their application to the languages of the world is indeed possible to achieve; the many publications devoted to this application bear evidence to this fact. The question then becomes, "What do these publications tell us?" "What is to be understood from examining the work of an analyst who constructs such theoretical classifications and defines a ‘type’?" In effect, the question becomes "What is
the allegory of this enterprise?"

6.5. The Allegory of 'Type' and 'Theoretical Classification'

It is proposed that what can be gained from an examination of the publications on these notions is an insight into the self of Western culture. The analyst approaches his rambunctious languages and linguistic data with both his theoretical and his pre-theoretical, or cultural, expectations. Any unified attempt to classify the world's languages (historically, typologically, or theoretically) will reflect the mind of the researcher; it will tell more about his own training and culture (and, perhaps, language) than it will about the world's languages. The more unified and internally consistent the classificatory proposal, the finer an example it will prove to be of that mode of thought which has been dominant in the West for so long (for further discussion, see L.A. Cummings 1986; and Elisabeth A. Cummings MS.).

And the languages of the world? Perhaps Sapir was right. By Horne's (1966) calculations of Sapir's typological proposals, there would almost be a separate 'type' for each language.

And people just go on talking.

6.6. The Anomaly of Object-Subject Languages

6.6.1. Responses to an Anomaly

Languages of OS sentential word order, then, stand out as anomalies, as oddities, among the languages of the world when considered from two separable (but often overlapping) linguistic approaches to sentential word
order. The Greenbergian approach finds them odd statistically because languages in which the S precedes the O overwhelmingly outnumber those with the inverted, OS order. The Prague school approach (and its derivatives) finds them odd (statistically, but also, one might say, psychologically) because what has been taken to be an emotionally or stylistically marked order of rheme-theme, appears to be the "basic", "neutral" order in these languages.

There are at least four possible responses to the uncomfortable question raised by OS languages. These will be presented, and either refuted or expanded upon.

6.6.2. Response 1: Status Quo

This response is that which would claim that OS languages are so statistically rare as to be insignificant in the grand scheme of "universal tendencies". These OS languages may exist; or they may not, only appearing in "surface order" to be of this odd configuration. If they exist, they are so insignificant statistically as to cast no blemish on the standard linguistic paradigm.

My refutation of this response will be of two parts. The first will be voiced by Pullum (1981), who compiled his catalogue of OS languages in order, he writes

to make it harder for linguists to go on believing that languages with object before subject are so marginal that only a tiny number of ill-fated human communities could tolerate them. I want to make it clear that a linguist who ignores the existence of the languages in question is ... acting like an anthropologist who decides to 'put aside' all the evidence of polygamy or matrilineality in order to be free to develop 'constructive' theories in which all cultures are monogamous and patrilineal. (1981:154)
The "patrilineal" and "monogamous" languages of linguists are those that conform to their theoretical and pre-theoretical expectations, and they search for a "dominant" word order. It is an exercise of cultural dominance which seeks to classify the languages of the world into a scheme conceived in a particular historico-cultural moment of the Indo-European West.

The second part of my refutation will take its analogy from science and mathematics, which ought to be respected sources for those who might make this first response; for they are most likely those who claim to be engaged in a scientific study of language.

Systems theorists define and model systems; they work in all fields of science and engineering. Linguists who work with formalisms and who conceptualize Language as a structure are also systems theorists. Until recently, modelizations devised by scientists and engineers have been purposely and admittedly "simple" systems. That is, it was known that there were external influences on a defined system other than those included in the model, but these influences were removed in order to be able to devise a system which was "linear". This linear system was solvable and understandable by mathematical tools; the relationships within the system were stable.

Now there is something different going on -- the study of nonlinear systems, also known as dynamics, chaos studies, and randomness studies. It is too long and complicated a story to tell here, and I refer the reader to Gleick (1987) for an introduction, and to West (1985). A brief quote will suffice to indicate the nature of this new kind of modelization:

The heart of the new snowflake model is the essence of chaos: a delicate balance between forces of stability and
forces of instability; a powerful interplay of forces on atomic scales and forces on everyday scales. Traditionally, because the surface tension effects [which are nonlinear, EAC] are so small, researchers assumed that for practical purposes they could disregard them. Not so. The tiniest scales proved crucial ... (Gleick 1987:310-311)

If linguistic modelization is ever to approach a description of Language it must cease being a "linear description", and enter into the study of random, chaotic, and dynamic systems.

This invocation of another big story has been made, remember, in order to aid my refutation of those language scientists who would dismiss OS languages from consideration because of their statistical "insignificance". Chaotic influences were initially shocking to scientists because tiny, seemingly insignificant differences of "input" suddenly were discovered to have very consequential effects on a system's behavior. Again, I quote from Gleick (1987:16-17), who describes an early recognition of this:

One day in the winter of 1961, wanting to examine one sequence at greater length, Lorenz took a shortcut. Instead of starting the whole run over, he started midway through. To give the machine its initial conditions, he typed the numbers straight from an earlier printout .... This new run should have exactly duplicated the old. Lorenz had copied the numbers into the machine himself. The program had not changed. Yet as he stared at the printout, Lorenz saw his weather diverging so rapidly from the pattern of the last run that, within just a few months, all resemblance had disappeared .... Suddenly he realized the truth. There had been no malfunction. The problem lay in the numbers he had typed. In the computer's memory, six decimal places were stored: .506127. On the printout, to save space, just three appeared: .506. Lorenz had entered the shorter, rounded-off numbers, assuming that the difference -- one part in a thousand -- was inconsequential. It was a reasonable assumption .... Yet in Lorenz's particular system of equations, small errors proved catastrophic .... Lorenz could have assumed something was wrong with his particular machine or his particular model .... But for reasons of mathematical intuition that his colleagues
would begin to understand only later, Lorenz felt a jolt: something was philosophically out of joint. The practical import could be staggering. (Gleick 1987:16-17)

Serious consideration of the anomaly of OS, rheme-theme languages must cause major re-consideration of the approach(es) and derivative modelizations which thought them not to exist.

This refutation may be put another way. Some have diagrammed individual languages to be individual "cognitive systems", complete with input and output. This is a model which is defined, linear, and deterministic. Using this system or model, the (sentential) output of the majority of the world's languages is SO, and theme-rheme in order. Now it is coming to light that a minority of languages have an inverse output: OS, and rheme-theme. One cannot argue with the data; they are there, and must be accounted for. This suggests, then, that something is missing, or wrong-headed, in either the definition of the system, or in the initial conditions of input to the system. Something is "out of joint". In terms of systems theorists, there are probably nonlinear, chaotic influences which have not been taken into account in the modelization. And, as Gleick explains, "Nonlinearity means that the act of playing the game has a way of changing the rules (1987:24). (Cf. Lyotard 1985). Godspeed to nonlinear linguistics!

6.6.3. Response 2: Compromise

Another possible response to the question raised by OS languages involves a compromise between data and theory. This response does not explicitly call for a re-consideration of the approaches which created the apparent anomaly but rather suggests an expansion of the typological approach.
For example, Payne's suggestion (already mentioned) that word order in some languages "... should be accounted for by syntactic factors, in others by pragmatic factors and in most languages by interacting factors of both types" (1987:802). This observation is in concurrence with studies by Mathesius (and others) of the Prague school, as noted by Uhlírova (1974:209) and Firbas (1974:12ff.). Uhlírova writes:

It was already pointed out by V. Mathesius that languages with rich morphology, such as Slavonic languages, usually have free word order and that word order is -- together with intonation -- the main means of expressing FSP [functional sentence perspective]. In contradistinction to the languages with fixed (grammaticalized) word order, such as English, it holds that the grammatical sentence elements in Slavonic languages are arranged one after another in utterances just according to their degrees of communicative dynamism, the importance of grammatical syntactic functions for the linear arrangement being very limited. Thus, it is not essential from the communicative point of view, whether, e.g., a Czech clause begins with a thematic, or rhematic element, or perhaps with a communicative element of another kind. (1974:209)

Mathesius was not making a compromise between theory and data, however; he was writing well before Greenberg's (1963) paper. Thus he was allowing his data to inform his observations. This does not deny the suggestion I now make that to state that some languages are more pragmatically motivated than others and that some languages are more grammatically motivated than others is perhaps somewhat odd. All languages have bits and pieces (that is, they involve a grammar) and all are spoken by people (that is, they involve pragmatic motivation). In the context of what was proposed in Chapter Three, a strict distinction between grammar and motivation or pragmatics is understood to be unnecessary (see Hardy 1988). The grammar, or workings, of a specific language permit its speakers to communicate their motivations.
To make a compromise between theory and data in the vein suggested by Payne is to let some languages be classified syntactically and others pragmatically and some as in-between. As has been indicated, this is an approach sympathetic with that of the Prague school; Vachek (1976), however, delineates the scopes of linguistic typological classification and linguistic characterology as being different enterprises.

6.6.4. Response 3: Revolution, or, Leave It

The "grammaticization" of Hopper (1987), referred to in Chapter Two, is, in part, a response to anomalies presented by close examination at the interface of linguistic grammatical theory and data from the world's languages (see also Hopper and Thompson 1980, 1984, and 1985). The incompatibility of OS languages, "free word order languages", and pragmatic order languages with a syntactically, sententially, and grammatically informed word order typology, again ought to cause reflection upon the manner in which "grammatical word order" languages have been studied. This manner of consideration, it was suggested in Chapters One and Two, is a result of the development of the discipline of linguistics and of the technologies of literacy.

This response has its philosophical and linguistic analogues as well, of course. It is the Heraclitean rather than the Aristotelian; the Bohemicists of Prague rather than the structuralists of France. It has the potential of orality rather than the potential of literacy, even in writing, paradoxically. Hopper's "repetitions" (1987:144), and Tyler's "meaning schemata" (1978:238-248) are not primarily motivated by a looking for stable, decomposable, classifiable structures, but by listening to people.
6.6.5. Response 4: Rhetoric and Culture, or, Take It and Run

This is a response not as radical as the last, but just as controversial, perhaps. This response, or path of inquiry, relies on the acceptance of the sentence as a semantic whole of some kind. Despite the frequent observation in this chapter that linguists have been "stuck inside sentences", it is clear that the sentence is some kind of semantic or communicative unit, as it is marked prosodically. The impatience with sentence linguistics partly results from awareness of two widely-held linguistic assumptions: first, that the sentence is some kind of primal unit; and second, that it is a unit that is processed in the brains of "language users" in terms of the encoding and decoding of constituent parts, or in terms of the application of some kind of innate rules. This response does not question the approaches which have led to the apparent anomaly of OS, rheme-theme, DIFFUSE-FOCUSED languages. Rather, it accepts and builds on the psychological difference implied by this apparent anomaly. With the coming to light of languages which have a sentential word order which is the grammatical, pragmatic, and psychological inverse of the majority of studied languages, one can postulate that experience is ordered differently for the speakers of these languages than it is for the speakers of SO, theme-rheme, FOCUSED-DIFFUSE languages.

This possibility rests on the belief that the order in which words are said not only mirrors their pragmatic importance within the discourse, but that this pragmatic "firstness" provides an entry into psychological or cognitive "firstness", or "salience". Thus, the pragmatic element which comes first in basic word order can be a guide to what is psychologically salient to the speakers of that language.
What is important here is the nature of this salience. It has been pointed out that, by the accounts of existing approaches, the majority of the world's languages are Subject-Theme-Topic initial in sentential order. Davis' term FOCUSED is analogous to these terms, but encompasses and accounts for more morphosyntax than these grammatical and pragmatic notions. The term FOCUSED implies a psychological attraction towards the discrete, the stable, the known. Pragmatically, it is aligned with TOPICS, semantically with what is KNOWN (in the sense of Chafe 1976). Grammatically, it is aligned with Subjects, definiteness, perfective aspect. Languages which are Subject-Theme-Topic initial in basic word order, then, find that what is psychologically FOCUSED most salient. That is, what is definite, known, stable, and discrete is more salient than what is indefinite, unknown, unstable and continuous. In Davis' terms, what is IMMEDIATE (rather than REMOTE) is most important to the speakers of SO languages, because it comes first in basic word order. This attention of focal awareness to what is IMMEDIATE is implemented in grammar. The psychological salience of definiteness, known-ness, stability and discreteness thus is continually implemented in the discourse of these languages; close attention to the syntactic, semantic and pragmatic strands of a discourse reflect this psychological salience.

Languages which are OS in basic word order, then, manifest a very different sense of what is psychologically salient. They involve sentence-initial placement of syntactic, grammatical and pragmatic elements which are the psychological opposite of what Davis calls FOCUSED; in his terminology, what is DIFFUSE comes initially. The term DIFFUSE aligns pragmatically with RHEME, semantically with what is NEW, grammatically with Objects,
indefinite reference, and imperfective aspect. Psychologically, focal awareness is attracted to what is diffuse or continuous rather than to what is focused and stable. Languages of OS basic word order, then, because they are "DIFFUSE-initial", manifest a psychological attraction to what is the inverse of "FOCUSED-initial" languages.

The idea that language somehow reflects something of the mind has been a persistent, tantalizing theme for centuries; in linguistics, the idea has had its ups and downs. This is not the place to tell that story, however. I will merely point out a few places where the idea has explicitly been stated in terms which are analogous to this response.

The first is from Prague functionalism, and their investigations of "theme", and "rHEME" in "communicative dynamism". Firbas (1974:11-12) notes Mathesius' awareness of a writing of Henri Weil (first published in 1844 in French); it seems that much of functional sentence perspective research (and its derivatives) is a working out and elaboration of Weil's suggestions. As Firbas notes (1974:22), "The basic distribution of CD [communicative dynamism; i.e., theme-rheme, EAC] would reflect what H. Weil has called the 'movement of the mind'". Weil's monograph was translated into English, entitled The Order of Words in the Ancient Languages Compared with that of the Modern Languages, and published in Boston in 1878. The languages in Weil's study, predictably for his era, and unfortunately for language universals research, were the following: Romance and Germanic, and Greek and Latin.

Greenberg, too, states this theme in the conclusion to his famous paper: "The order of elements in language parallels that in physical experience or the
order of knowledge" (1963:81). In a much more recent paper (1985), Greenberg explores in some detail this theme with reference to "place, time, and discourse deixis". This paper appears in a volume entitled Iconicity in Syntax (Haiman 1985); "iconicity" is a term from Haiman (1980), and is serving as a rallying point for investigation of the old, tantalizing theme under discussion here.

It is interesting to note that that the evidence from OS languages is drawing some attention by those engaged in studies of "iconicity". For example, Givon (in the same volume), after referring to Malagasy (VOS) and to Hixkaryana (OVS), writes the following:

What we have here, presumably, is two different ways human languages may resolve the potential conflict between a figure-ground, perceptual-saliency oriented semantic iconicity principle, and a task-urgency oriented pragmatic iconicity principle. And while the bulk of human languages resolve the conflict by going along with the more primal semantic principle (SVO, SOV, VSO), a small group -- VOS, OVS -- chose to go with the clearly more complex and less obvious pragmatic principle. (1985:212)

This response, or path of inquiry, may possibly be taken further. If there is such a psychological difference as implied by these language studies, then this attraction to the continuous will be manifest in the means of aesthetic expression available to a culture. But that is another big, tantalizing story.
CHAPTER SEVEN

CONCLUSION

7.1. Overt Considerations, or, Summary of Chapters

7.1.1. First Chapter

This chapter examined the discourse of linguistics from a number of stances. First, it rendered a caveat on the study of language when conducted with scientific and logico-mathematical expectations. It questioned whether the nature of formal models permit any contribution to the understanding of languages, and asks whether such formalizations are not, rather, a mirror of the expectations of those who devise them. These expectations were seen to be engendered first, by Euro-American culture (in the guise of alphabetic literacy, scientific methodology, and a logico-mathematical mode of thought), and second, by Indo-European grammar. Linguistics, or the twentieth-century discourse on language, when practiced on this basis, was seen to richly exemplify the representational mode of thought. It was observed that this discourse contributes to the "fragmentation of knowledge", and thus represents a symptom of a wider dis-ease in Western culture.

7.1.2. Second Chapter

The discussion then moved to a consideration of linguistics as a classificatory discourse. Though typological classification became a focus, study was made of the historical, methodological, and epistemological interplay between this "strand" of linguistics, and the two other traditional strands: historical linguistics ("genealogical classification") and theoretical linguistics.
("theoretical classification"). All three strands were discussed in detail as constructed systems of classification seeking to impose order on, or to make universal statements about, the plethora and diversity of the languages of the world. Notice was made of the power and dominance in linguistic thought of the Platonic metaphor.

Within the study of language, the progressive detachment of talk from people was commented upon. The simultaneous appearance of the autonomous discipline of linguistics was noted, and the caveat of the previous chapter was invoked and expanded upon.

A concentration on typological classification was introduced by a historical sketch of nineteenth-century typological proposals, and continued by a more in-depth discussion of twentieth-century proposals. Particular consideration was given to the thought and writing of Edward Sapir and Joseph Greenberg.

The chapter concluded with an invocation of the writings of contemporary linguists from each of the three traditional strands, or subfields, of linguistics. In these writings there was seen to be a questioning and lack of acceptance of the manner in which the study of language(s) has been conducted. It was observed that the classificatory discourse of linguistics is in the process of dissolving.

7.1.3. Third Chapter

The focus on typological classification was returned to in this chapter by a recalling of the influence of Joseph Greenberg’s proposals of a syntactic typology based on word order.
The motivation for the following two chapters was introduced by a discussion of the statistical rarity, by the account of Greenbergian typology, of that language type in which the Object precedes the Subject in basic word order. The semantic approach to language(s) taken in the language interpretations to follow was then introduced. It was suggested that a more comprehensive and interesting understanding of word order than that of syntactic typology might be had by considerations of a semantic, pragmatic, functional, and psychological nature, in combination with a close attention to the data of individual languages.

The manner of presentation of the two languages to follow was discussed as being the telling of a story rather than as the reporting of a scientific hypothesis. It was pointed out that the understanding of how languages work to be employed in the next two chapters was incompatible with linguistic modelling or formalization.

7.1.4. Fourth Chapter

This chapter provided a detailed study of word order in Hixkaryana (Carib, OVS). The initial interest was to determine why the Object preceded the Subject in basic word order. First, the arguments and evidence for the language being OVS in syntactic order were presented. Then a close examination of word order possibilities in Hixkaryana was embarked upon, from the semantic-psychological perspective introduced in Chapter Three.

Hixkaryana was found to be consistently RHEME-initial in its sentential order. (Direct) Objects in the language were seen to be consistently tied to the semantic REMOTENess characteristic of RHEME. Thus, when nominally
expressed, Objects most often occurred sentence-initially.

Whether a sentential element was a grammatical Subject, Object, or Verb, was seen not to bear much significance in an understanding of Hixkaryana discourse.

7.1.5. Fifth Chapter

An explanation and interpretation of Tzotzil (Mayan, VOS) word order constituted the theme of this chapter. Note was made of the extent of VOS word order within the Mayan family. Reference was made to the problems encountered by researchers on Mayan languages in the determination of basic word order, and to discourse studies of various languages which have proved more elucidating in their discussions of word order.

The semantic-psychological factors determinant in Tzotzil for an explanation of its Object-Subject word order were found to be that the language expresses sentence-initially that element which is most cognitively "active" or DIFFUSE, and that what is most cognitively "stable" or FOCUSED is expressed in sentence-final position. Once again, whether an element was a grammatical Subject or Object was seen to have limited relevance for an understanding of word order in the language.

In the concluding remarks on Hixkaryana and Tzotzil, it was observed that not only grammatical syntactic elements seemed to be relatively unimportant for an understanding of sentential order, but it was intimated that discourse in these languages is not structured on the sentential level at all. This observation has implications not only for the tradition of typological
classification, but for the tradition of theoretical classification as well.

7.1.6. Sixth Chapter

The discussion in this chapter returned to the broad considerations of the first two chapters. The implications of the language interpretations of Hixkaryana and Tzotzil for the classificatory approaches to language were discussed.

The focus on typological classification was concluded, first by a discussion of the concept of "markedness", which has enabled this classification. Serious philosophical, commonsensical, and methodological problems with this concept were noted. Second, four distinct assumptions upon which word order typology is based were presented; all were found to be untenable.

It was suggested that what can be mostly gleaned from study of the many publications devoted to the classificatory discourse of linguistics is an understanding of the expectations and minds of their writers. The allegory of the classificatory approaches to the world's languages, it was suggested, is that they provide an insight into the self of Western culture, and into a mode of thought which has been long dominant in that culture. The dissolving of classificatory linguistics is but one of the encouraging rumblings abroad that both that culture and that mode of thought are not what they were once thought to be.

The Object-Subject languages interpreted in Chapters Four and Five were seen to represent anomalies, not only for the tradition of word order typology, being of this "rare" word order, but also for the theoretical approaches based
on functional-pragmatic and psychological considerations. Their word order was pointed out as being grammatically Object-Subject, functionally rheme-theme, semantically RHEME-TOPIC, and psychologically DIFFUSE-FOCUSED. All of these orders are "inverted" when considered from these viewpoints.

Four possible responses to anomaly represented by these languages were offered. Each of them was either refuted or left open. Those left open constitute alternatives to the study of language(s) in the face of the dissolution of the classificatory approaches.

7.2. Covert Considerations Revisited, or, Now to Begin

Thus, the classificatory discourse of linguistics constitutes a valiant effort towards the enunciation of the universal statements spoken of in the introduction to the present work. To the Greek question of "the one and the many" this discourse has concentrated on "the one". In the manifestation of the question within linguistics as "universality and diversity", the utilization of Aristotelian classification as a means of "talking about talk" has necessitated pre-occupation with universality.

Linguists, as they appropriate the languages of the world, have long professed to be democratic rather than prescriptive in their attitude towards these languages. The study of the imagination and discourse of linguistics in the present work, however, suggests that the means of this appropriation is culturally specific to a Graeco-Latin, literate, scientific West, grammatically specific to a Graeco-Latin interpretation of Indo-European languages, and linguistically specific to Indo-European. If this is accepted, then the seeking of
the articulation the universal statements concerning language, languages, and speakers are not democratic at all, but rather are despotic moves of cultural and linguistic dominance.

The strong impulse within classificatory linguistics which has sought to see the Platonic ideal form, and to speak it, has also resulted in an approach to languages which has caused them to be de-peopled, de-voiced, and de-storied. The dissolving of the classificatory approaches, it is hoped, will enable the talk to put back into talking about talk.
BIBLIOGRAPHY


Bach, Emmon, and Robert T. Harms, eds. 1968. Universals in Linguistic


New Orleans: Middle American Research Institute, Tulane University
Occasional Paper 1.


Brody, Jill. 1984b. "Some Problems With the Concept of Basic Word Order".
*Linguistics* 22, 711-736.

Brown, Penelope. 1979. *Language, Interaction, and Sex Role in a Mayan
Community*: A Study of Politeness and the Position of Women. Doctoral
Dissertation, University of California, Berkeley.

Campbell, Lyle, and Marianne Mithun, eds. 1979. *The Languages of Native
America: Historical and Comparative Assessment*. Austin-London:
University of Texas Press.

Campbell, Lyle. 1979. "Middle American Languages". In Campbell and
Mithun (1979), 902-1000.

Are We Now?" *Annual Review of Anthropology* 14, 187-198.


Casagrande, Joseph B. 1963. "Language Universals in Anthropological
Perspective". In Greenberg (1963a), 220-235.

Chafe, Wallace L. 1976. "Givenness, Contrastiveness, Definiteness, Subjects,
Topics, and Point of View". In Li (1976), 25-55.

Press.

Cole, Michael and Sylvia Scribner. 1974. *Culture and Thought: A


Stockholm: Almquist and Wiksell.


Chicago: The University of Chicago Press.

Conte, Maria-Elisabeth von, Anna Giacalone Ramat and Paolo Ramat, eds.
Tubingen: Niemeyer.


Cummings, Elisabeth A. 1988. MS. "Some Room to Talk and Play: Three Postmodern Motifs".


Davis, Philip W. 1987. MS.c. Language and Intelligence.


Derbyshire, Desmond C. 1977b. "Discourse Redundancy in Hixkaryana". IJAL 43, 176-188.


Languages". *IJAL* 4, 192-214.


Macmillan.


Language and Society 5, 67-80.


Lehmann (1978a), 169-222.


Mathesius, Vilem. 1928. "On Linguistic Characterology with Illustrations


Pinkerton, Sandra. 1976b. "Ergativity and Word Order". In Pinkerton


Schachter, Paul. 1976. "The Subject in Philippine Languages: Topic, Actor, Actor-Topic, or None of the Above". In Li (1976), 491-518.


Published under the auspices of CIPL (The Hague), Tokyo.


Technology Press of MIT.

