NOTE TO USERS

This reproduction is the best copy available.

UMI®
ISSUES IN COMPARATIVE UTO-AZTECAN MORPHOSYNTAX

by

Jason Derek Haugen

Copyright © Jason Derek Haugen 2004

A Dissertation Submitted to the Faculty of the

DEPARTMENTS OF ANTHROPOLOGY AND LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2004
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Jason Derek Haugen entitled Issues in Comparative Uto-Aztecan Morphosyntax and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Dr. Heidi Harley

Dr. Jane H. Hill

Dr. Terence Langendoen

Dr. Andrew Carnie

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director: Dr. Heidi Harley

Dr. Jane H. Hill
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: [Signature]
ACKNOWLEDGMENTS

This dissertation could not have been written without the support of many people during my graduate school years in Tucson. First, of course, was my committee, who could not have been more supportive, inspirational, and generally fantastic. Any student would have been lucky to have had any one of these amazing individuals to consult with, on all matters, on a regular basis. Thanks to Heidi Harley, Jane Hill, Terry Langendoen, and Andrew Carnie. I also give many thanks to Eloise Jelinek. Although she was not technically on my dissertation committee, the ideas contained herein are extensively permeated by her work. It was Eloise, in her summer 1999 pedagogical workshop on the structure of Yaqui, who introduced me to the wonderful world of Uto-Aztecan linguistics. She has set a high standard for all linguists to follow when pursuing the study of indigenous North American languages at the descriptive, theoretical, and applied levels. She provided helpful discussion of many of the ideas presented in the following pages. In addition, I must also thank Maria and Rosario Amarillas for many interesting and very informative hours discussing the Yaqui language. All of the above-mentioned individuals have had a concrete and highly positive impact on the direction and final product of this research, but any errors or other infelicities that might remain are mine alone.

My educational experience in the Joint PhD Program in Anthropology and Linguistics has been a diverse one, and I would like to thank all of the faculty with whom I have taken courses over the years: in addition to my committee members, in Linguistics I would like to thank Diana Archangeli, Andrew Barss, Dick Demers, Sheila Dooley-Collberg, Mike Hammond, Massimo Piatelli-Palmarini, Natasha Warner, Mary Willie, and Elizabeth Winkler; and in Anthropology I thank Norma Mendoza-Denton, Susan Philips, Rudi Gaudio, Linda Green, and William Longacre. I had the good fortune of attending the 1999 American Indian Language Development Institute, and I would like to thank my instructors Lucille Watahomigie and the incomparable Ken Hale. I was also lucky to have had the opportunity to attend the 2001 LSA Summer Linguistics Institute at UC-Santa Barbara, for which I would like to single out Emmon Bach, Paul Chapin, and Marianne Mithun, to thank for their highly informative courses. The idea of pursuing graduate study in Anthropology and Linguistics was inspired by my undergraduate course-work at the University of Texas at Austin. I thank in particular Mark Louden, Joel Sherzer, Carlota Smith, and Keith Walters, for their support and encouragement in those early years.

I also want to thank all of the Uto-Aztecanists that have shared their knowledge and opinions with me over the years; included in this group are to be noted especially Dirk Elzinga, Colleen Fitzgerald, Larry Hagberg, Ken Hill, Dave Shaul, Constantino Martínez Fabian, Gabriela Caballero, Ana Aurora Medina Murillo, and Ofelia Zepeda. Other UA linguists who have had a positive influence on me include Rudy Troike, as well as my graduate student elders, especially Sonya Bird, Debbie Cole, Sean Hendricks, Will Lewis, Jessica Maye, Mizuki Miyashita, Laura Moll, Erin O’Bryan, Meghan O’Donnell, Kazutoshi Ohno, and Jessica Weinberg.
I can never give enough thanks to my graduate student contemporaries here at Arizona, with whom I shared many hours of struggle and triumph, and under both of these conditions, fun. Foremost among my colleagues I must point to Cathy Hicks Kennard, with whom I had the great fortune of making the transition from UT undergrad to UA grad, and along the way I have also been very privileged to get to know Andy and Christopher as well. The Hicks Kennards are like a second family to me, and I cannot imagine having gone through this experience without their constant friendship. I also thank the other fellows of my cohort: Ana Carla Bruno, Bob Kennedy, Sarah Longstaff, Peter Norquest, and Katie Russell, each of whom played an instrumental role in helping me run that initial gauntlet, in addition to being there when needed well after those frenzied first semesters. Plenty of other friends and colleagues have also come along since then, for whom I would like to give special thanks to William Alexander, Luis Barragan, Ruby Basham, Phil Cash Cash, Tania Granadillo, Rachel Hayes, Travis Louthain, Scott Jackson, Emily Kidder, Charles Lin, Heidi Orcutt, Jianyuan Wang, Adam Baker, Lynnika Butler, Polly LeBlanc, Jaime Parchment, Peter Richtsmeier, Dan Siddiqi, Ben Tucker, and all the others that I have gotten to know over the years.

My graduate career was financially supported by several institutions. Thanks to the US Department of Education for a Jacob K. Javits Fellowship, and the Graduate and Professional Student Council at the UA and the Social and Behavioral Sciences Research Institute at the UA for grants which have funded various aspects of this research. And many thanks also to the Departments of Anthropology and Linguistics for their financial support. Thanks to the department heads at UA over the years—Tom Bever and Mike Hammond in Linguistics, and William Longacre and John Olsen in Anthropology. I also thank the staffs of these departments, especially Rosemary Emery and Gwen Weber.

Thanks to all of the various members of the Pima Street Hockey League over years, as well as to the members of the UA Co-Rec Floor Hockey Teams: the Men’s League team (Co-Rec Champs 2001, 2002), and the Co-Ed team (Co-Rec Champs 2001, 2003, 2004), as well as my racquetball partner/nemesis Evan Grant; all in addition to my other friends in Tucson, Arizona.

No dissertator should ever do their work without the unquestioning support of a loving pet; I have had three of the best—thanks to Dominique and Goldie for always being around, and thanks to the late Hobo, with whom many hours of leisure were spent relaxing from the researching and writing of this dissertation. He is still greatly missed.

I also thank my family for everything that they have done for me, and for being understanding in my extended absences from them. Thanks to my mother and father, Lisa, Scott and the boys, and to everybody else. Last, but far from least, I thank Amy Vlassia Margaris, whom I met here in grad school, and who has been there for me in the best times and the worst, and whose loving support has made everything possible and worthwhile.
# TABLE OF CONTENTS

LIST OF TABLES ................................................................................................. 10
LIST OF FIGURES ................................................................................................. 11
ABSTRACT ........................................................................................................... 12

## PART I: INTRODUCTION AND BACKGROUND

### CHAPTER 1: INTRODUCTION ........................................................................... 14

### CHAPTER 2: THE UTO-AZTECAN LANGUAGE FAMILY

2.1. The Uto-Aztecan languages and their classification ........................................ 20
2.2. Uto-Aztecan grammar ................................................................................... 25
  2.2.1. Overview of Uto-Aztecan syntax ............................................................... 27
  2.2.2. DPs and case-marking ............................................................................ 30
  2.2.3. Subject and object clitics ....................................................................... 34
2.3. Uto-Aztecan historical linguistics and cultural prehistory ............................ 42
2.4. Uto-Aztecan and linguistic theory ................................................................. 46

### CHAPTER 3: LINGUISTIC THEORY AND HISTORICAL (MORPHO)SYNTAX

3.1. Introduction .................................................................................................. 49
3.2. General assumptions about the nature of language and how we should study it .............................................................................................................. 49
3.3. The Pervasive Syntax Perspective and Distributed Morphology .................. 53
3.4. Comparative syntax and syntactic reconstruction ......................................... 60
  3.4.1. Grammaticalization and Grammaticalization Theory ............................ 60
  3.4.2. Grammaticalization and Minimalism—Roberts and Rousseau (2003) ...... 64
  3.4.3. On syntactic reconstruction—Harris and Campbell (1995) ....................... 67
3.5. Summary ....................................................................................................... 73
# TABLE OF CONTENTS - Continued

## PART II: PROSODIC MORPHOLOGY

### CHAPTER 4: PROSODIC MORPHOLOGY AND CONSTRAINT-RANKING IN UTO-AZTECAN HISTORICAL PHONOLOGY

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Introduction: Raimy's Modular-Derivational approach to reduplication</td>
<td>74</td>
</tr>
<tr>
<td>4.2. Reduplication in Uto-Aztecan</td>
<td>77</td>
</tr>
<tr>
<td>4.2.1. &quot;Marked heavy syllable reduplication&quot;</td>
<td>81</td>
</tr>
<tr>
<td>4.2.1.1. Yaqui (SUA)</td>
<td>82</td>
</tr>
<tr>
<td>4.2.1.2. Mayo (SUA)</td>
<td>86</td>
</tr>
<tr>
<td>4.2.1.3. Guarijio (SUA)</td>
<td>87</td>
</tr>
<tr>
<td>4.2.1.4. Nahuatl (SUA)</td>
<td>89</td>
</tr>
<tr>
<td>4.2.1.5. Numic (NUA)</td>
<td>90</td>
</tr>
<tr>
<td>4.2.2. Mora affixation</td>
<td>92</td>
</tr>
<tr>
<td>4.2.2.1. Mora affixation in Yaqui (SUA)</td>
<td>92</td>
</tr>
<tr>
<td>4.2.2.2. Mora affixation in Tepecano (SUA)</td>
<td>93</td>
</tr>
<tr>
<td>4.2.2.3. Mora affixation in the Numic languages (NUA)</td>
<td>95</td>
</tr>
<tr>
<td>4.3. Capturing the Uto-Aztecan generalizations through OT constraint-ranking</td>
<td>99</td>
</tr>
<tr>
<td>4.4. Some complicating data: A three-way distinction in Tohono O'odham</td>
<td>105</td>
</tr>
<tr>
<td>4.5. Against the alternative analysis</td>
<td>113</td>
</tr>
<tr>
<td>4.5.1. Reduplication and prosodic units</td>
<td>114</td>
</tr>
<tr>
<td>4.5.2. Constraint-ranking, or no?</td>
<td>117</td>
</tr>
<tr>
<td>4.5.3. On the morphological status of the reduplicant</td>
<td>121</td>
</tr>
<tr>
<td>4.6. Conclusion</td>
<td>123</td>
</tr>
</tbody>
</table>

### CHAPTER 5: ON REDUPLICATION AND SYNTAX

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1. Introduction</td>
<td>125</td>
</tr>
<tr>
<td>5.2. Travis on the syntax of reduplication</td>
<td>125</td>
</tr>
<tr>
<td>5.2.1. Travis (1999): &quot;A syntactician's view of reduplication&quot;</td>
<td>126</td>
</tr>
<tr>
<td>5.2.2. Travis (2001): &quot;The syntax of reduplication&quot;</td>
<td>133</td>
</tr>
<tr>
<td>5.3. On the syntax of &quot;phonological reduplication&quot;</td>
<td>139</td>
</tr>
<tr>
<td>5.4. Conclusion</td>
<td>150</td>
</tr>
</tbody>
</table>

## PART III: DERIVATIONAL MORPHOLOGY

### CHAPTER 6: NOUN INCORPORATION AND DENOMINAL VERBS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1. Introduction</td>
<td>151</td>
</tr>
<tr>
<td>6.2. Denominal verbs and incorporation, or denominal verbs as incorporation?</td>
<td>156</td>
</tr>
</tbody>
</table>
### TABLE OF CONTENTS - Continued

6.2.2. Mithun (1984) on the (non-)syntactic status of NI..............................165
6.2.4. Mithun (1986)—Rejoinder to Sadock..................................................170
6.3. Denominal and noun incorporating verbs in Hopi: Evidence for a unified 
analysis........................................175
   6.3.1. Noun incorporation in Hopi..........................................................177
   6.3.2. Denominal verbs in Hopi..............................................................179
   6.3.3. Hopi denominal verbs as incorporating verbs..................................182
6.4. Against a Lexicalist account of noun incorporation: Rosen (1989)............184
6.5. Summary...............................................................................................193

### CHAPTER 7: NOUN INCORPORATION IN UTO-AZTECAN

7.1. Introduction............................................................................................195
7.2. Categories of Noun Incorporation in Uto-Aztecan: Overview.....................196
   7.2.1. N-V Compounding.................................................................196
   7.2.2. Syntactic NI.........................................................................200
   7.2.3. “Object polysynthesis”............................................................203
   7.2.4. Classificatory NI.................................................................205
7.3. Noun incorporation in Uto-Aztecan I: Results of the survey.......................207
   7.3.1. Getting over the distinction between derivational morphology and NI...208
   7.3.2. A typology for Uto-Aztecan noun incorporation...............................210
   7.3.3. Recasting Mithun (1984)’s implicational hierarchy of NI types...........210
   7.3.4. Further issues in Uto-Aztecan NI.................................................211
      7.3.4.1. Instrumental prefixes......................................................211
      7.3.4.2. Subject incorporation......................................................215
7.4. Noun incorporation in Uto-Aztecan II: The data....................................220
   7.4.1. Numic: Comanche.................................................................220
   7.4.2. Takic: Cupeño.............................................................................228
   7.4.3. NUA Isolate: Hopi.................................................................233
   7.4.4. Tepiman: Tohono O’odham.......................................................235
   7.4.5. Taracahitic: Yaqui.................................................................240
   7.4.6. Corachol-Aztec: Nahuatl.........................................................250
   7.4.7. Denominal verbs across Uto-Aztecan..........................................250
7.5. Conclusion..............................................................................................251

### CHAPTER 8: ON THE SYNTAX OF NOUN INCORPORATION: 
INTEGRATION AND CONFLATION, MOVE AND MERGE

8.1. Introduction............................................................................................253
8.2. Theoretical background: Incorporation and conflation.............................254
# TABLE OF CONTENTS - Continued

8.2.2. Hale and Keyser (2002): Denominal verbs and conflation..............................261
8.3. Late Insertion as the solution to the hyponymous object problem......................266
8.4. Non-DP complements to V: N-V compounding revisited........................................278
8.4.1. N as complement to V: A formal account of N-V compounding.........................279
8.4.2. Other complements to V......................................................................................281
8.5. Object polysynthesis and pronominal arguments....................................................283
8.6. Conflation and non-object noun incorporation in Uto-Aztecan.............................287
8.6.1. Instrumental prefixes and manner conflation.......................................................288
8.6.1.1. Instrumental prefixes in Northern Paiute.........................................................290
8.6.1.2. A syntactic analysis of instrumental prefix constructions.................................295
8.6.2. Subject incorporation..........................................................................................299
8.6.2.1. Unaccusative incorporation..............................................................................301
8.6.2.2. N-V compounding with "agentive" interpretation...........................................304
8.7. Conclusion..............................................................................................................308

## PART IV: CHANGE IN MORPHOLOGICAL TYPE

### CHAPTER 9: ON THE GRADUAL DEVELOPMENT OF POLYSYNTHESIS IN NAHUATL

9.1. Introduction.............................................................................................................309
9.2. Perspectives on polysynthesis................................................................................311
9.3. Subject and object polysynthesis in cross-linguistic perspective .........................319
9.3.1. Pronominal order in Athabaskan.........................................................................323
9.3.2. Pronominal order in Cupéño..............................................................................325
9.4. Structure of Nahuatl..............................................................................................328
9.5. Nahuatl in Uto-Aztecan context.............................................................................332
9.5.1. Polypersonalism..................................................................................................332
9.5.1.1. Object pronominals.........................................................................................333
9.5.1.2. Subject pronominals.......................................................................................335
9.5.2. Syntactic noun incorporation..............................................................................335
9.5.3. Related issues.....................................................................................................337
9.5.3.1. Discourse reference for incorporated nouns...................................................337
9.5.3.2. Discontinuous constituents.............................................................................338
9.5.4. Summary............................................................................................................339
9.6. Conclusion..............................................................................................................340

REFERENCES...............................................................................................................341
LIST OF TABLES

Table 2.1. Proto-Uto-Aztecan consonants..........................................................26
Table 2.2. Proto-Uto-Aztecan vowels.................................................................26
Table 2.3. Subject clitics across Uto-Aztecan...................................................37
Table 4.1. Markedness typology for Uto-Aztecan heavy syllable reduplication........100
Table 6.1. Cluster of properties associated with noun incorporation.................191
Table 7.1. Uto-Aztecan NI types and their characteristics.............................210
Table 9.1. Grammaticalization trajectory for (object) polysynthesis..................336
LIST OF FIGURES

Figure 2.2. The Uto-Aztecan languages & major sub-groups...............................21
Figure 2.2. Map of the Uto-Aztecan languages.................................................22
Figure 3.1. The DM Model.....................................................................................57
This dissertation seeks to test recent important theoretical ideas in the Principles and Parameters and Distributed Morphology frameworks against data from the relatively under-studied Uto-Aztecan languages. In this work I focus on the morphology of reduplication, noun incorporation and related derivational morphology, and the diachronic development of the polysynthetic morphological type in one sub-branch of the family (Corachol-Aztecan).

With respect to prosodic morphology, I argue that the comparative Uto-Aztecan evidence suggests that reduplicants should be viewed as morphological pieces, and I analyze them as Vocabulary Items inserted into syntactic slots at Morphological Structure. I also argue that the evidence of cognate reduplication patterns across Uto-Aztecan supports a prosodic view of morphology, as well as the constraint-ranking approach to morphophonology.

With respect to noun incorporation and derivational morphology, I argue that the comparative Uto-Aztecan evidence supports the view that denominal verbs are a subclass of noun-incorporating verbs. I survey the noun incorporation types in Uto-Aztecan and classify NI in these languages into four types: N-V compounding, syntactic NI, classificatory NI, and “object polysynthesis”. I offer a unified syntactic account of these types, maintaining that each is formed via head-movement in syntax. I provide a novel approach to hyponymous objects, suggesting that these are in argument positions, and that they are derived via the Late Insertion of material that is not cognate to the incorporated noun, but which is inserted into the lower copy of a movement chain. Non-
theme "nominal" roots incorporated into verbs, such as instrumental prefixes, are analyzed as adverbial elements merged directly into the verbal position.

Finally, I argue that this theoretical analysis of NI leads naturally to a diachronic account of the development of polysynthesis in Nahuatl. I show that the crucial aspects of polysynthesis, subject and object pronominal marking on the verb as well as syntactic noun incorporation, have analogues elsewhere in Uto-Aztecan, and I offer a reconstruction of the likely stages of the development of polysynthesis in Nahuatl, each of which have attestation elsewhere in the family.
CHAPTER 1
INTRODUCTION

The purpose of this dissertation is to test important recent theoretical claims within the Principles and Parameters approach to morphology and syntax against data from the relatively under-studied Uto-Aztecan language family. Any theoretical approach to cross-linguistic variation must be tested against a wide variety of typologically distinct languages in order that the full range of linguistic phenomena are accounted for, and in order to confirm that things that our models predict should not happen in natural language actually do not happen. In the ideal case, if a theoretical model is to be taken seriously, such investigation should also have implications for the study of language change—in the case of syntax, this investigation should be important for the delineation of the principles and parameters of Universal Grammar, and the conditions on the setting of parameters.

It is the goal of this dissertation to address these questions by taking a close look at the comparative morphosyntax of the Uto-Aztecan languages. To my knowledge this is the first effort at testing claims within the Principles and Parameters framework using strictly comparative Uto-Aztecan data, although certain prominent analyses have used data from Uto-Aztecan languages to make particular theoretical points.

This dissertation is divided into four parts.

Part I, Introduction and Background, is divided into two chapters, in addition to the present one. In Chapter 2 I provide the background to Uto-Aztecan historical linguistics necessary to understand the comparative aspects of later chapters. Chapter 3 presents the
theoretical model and assumptions under which later analyses are presented—Distributed Morphology (DM). It is here that I introduce the Pervasive Syntax Perspective (PSP), the organizing hypothesis that drives later discussions.

Part II, *Prosodic Morphology*, begins the theoretical investigation into comparative Uto-Aztecan morphosyntax by looking at the morphology of reduplication in the context of Uto-Aztecan historical phonology. Prosodic morphology is an area of research crucial to any piece-based theory of morphology, because it is not at all obvious how what appears to be a process of altering prosodic structure can constitute a "piece" of morphology. Within this theme we can include the well-known phenomenon of reduplication (and morphological mora augmentation), but also truncation (subtractive morphology), ablaut, morphological metathesis, etc., all of which have been discussed within the framework of another piece-based theory of morphology: Combinatorial Morphology (Stonham 1994). Of these processes, only the first two are prevalently active in Uto-Aztecan, morphological truncation being particularly developed in Tepiman. However, I focus in Chapter 4 only on reduplication (including mora affixation), since it is ubiquitous in the family, and since most Uto-Aztecan languages actually employ multiple reduplicative morphemes for a variety of functions, including the modification of suffixes. Thus, Uto-Aztecan languages provide us with a perfect opportunity to explore the interaction of reduplication with other aspects of the grammars of very closely related languages.

Chapter 4 discusses previous work on reduplication within the piece-based theory of Distributed Morphology, the modular-derivational theory of reduplication proposed by
Raimy (2000), and I show that three crucial aspects of Raimy’s model are undesirable in accounting for the cross-linguistic facts of Uto-Aztecan reduplication. Contra Raimy, I argue that Uto-Aztecan reduplication is best accounted for by reference to prosodic organization in phonology, a constraint-based, rather than derivational, approach to generating phonological outputs, and the existence of reduplicants as morphemes rather than as epiphenomena of other phonological processes.

Chapter 5 discusses the syntax of Uto-Aztecan reduplication, and here I show that an Optimality Theoretic model is not necessarily inconsistent with an approach to morphology that builds morphological structure in syntax. I argue that Morphological Structure is the input to Phonological Structure, which I construe as the input to OT-style constraint-evaluation.

Part III, Derivation in Uto-Aztecan Syntax, focuses on a central area of recent theorizing in morphological theory: the nature of noun incorporation and related derivational processes, and their relationship to syntax. Much recent theoretical work in morphology has focused on the role of syntax in creating morphological conglomerations that may be regarded, in a pre-theoretical way, as “complex words”. In some cases theorists have posited that complex words are formed via the movement of syntactic heads, or by means of other syntactic processes (e.g. Baker 1988, Sadock 1991, Julien 2002, among many others,), although there are also other prominent alternative views.

Given the PSP adopted in Chapter 3, in this part of the dissertation I focus on those processes that form verbs from what are ostensibly noun roots: noun incorporation (of several distinct types) and denominal verb formation. I observe that these processes are
prevalent in the Uto-Aztecan languages, but that in some cases there are two distinct syntactic mechanisms that lead to the formation of various kinds of derived verbs. Some involve movement, while others involve Merge alone (i.e. base-generation), and I discuss the formal criteria necessary to distinguish between the two.

Part III is organized as follows. Chapter 6 reviews the previous literature on noun incorporation (NI) and denominal verb formation, and argues, using empirical evidence from Hopi, that the two processes are derivational in nature, and that these constructions are formed by identical syntactic mechanisms. In addition, in this chapter I critique a prominent Lexicalist approach to NI phenomena (Rosen 1989). In Chapter 7, I provide an overview of NI phenomena found in the Uto-Aztecan languages, and present four heuristic categories useful for comparing NI across these languages: N-V compounding, syntactic NI, classificatory NI, and “object polysynthesis”. In Chapter 8, I discuss the nature of the syntactic mechanisms that derive each of the NI types. Here I argue that each of the four NI types can be unified under a head movement account of incorporation (cf. Baker 1988), for those incorporated nouns that have a theme theta role. Other noun incorporation-like constructions, such as instrumental prefixes, are analyzed as compounds brought about through Merge. In this discussion I define and elaborate the crucial notions of incorporation and conflation (Hale and Keyser 2002), and I offer a novel approach to hyponomous arguments that makes use of the notion of Late Insertion of different lexical material into the head and lower copies of a movement chain.

The area of research under focus in this section lies at the heart of the interface of syntax and morphology in current syntactic theorizing, and these areas provide an
empirical testing-ground on which the Lexicalist Hypothesis may stand or fall.

Following the PSP, I argue that all of these processes must occur in syntax proper. This position entails the reconstructability of particular derivational processes insofar as we can reconstruct particular derivational morphemes.

Finally, in Part IV, Change in Morphological Type, I extend the synchronic analysis presented in Chapter 8 to the question of the historical development of NI in Uto-Aztecan. In Chapter 9, On the Gradual Development of Polysynthesis in Nahuatl, I argue that the comparative evidence from Uto-Aztecan suggests that we can see "precursors" to the obligatory subject and object marking on verbs that is unique to Nahuatl (or Corachol-Aztecan) within the Uto-Aztecan family. I argue that the crucial phenomena of subject and object clitics, in addition to syntactic noun incorporation, all lead to a plausible series of stages in the historical development of a new morphological type (i.e. polysynthetic) in the Uto-Aztecan family.

This dissertation was written with two distinct (although not necessarily mutually exclusive) audiences in mind. First, several of the chapters are oriented toward the community of linguists who are more interested in description than theory. In particular, Chapters 4 and 7 focus on presenting a wide range of data involving Uto-Aztecan reduplication patterns and noun incorporation constructions, respectively. However, I also hope that the theoretical orientation(s) that I have adopted elucidate the patterns that we see in crucial ways. With respect to reduplication, the constraint-ranking approach leads to cross-linguistic generalizations that would be hard to capture using the traditional tools of historical phonology. Similarly, the theoretical analysis presented in Chapter 8
unifies the four disparate categories of NI presented in Chapter 7, based on the assumptions adopted under the tenets of the PSP. In Chapter 9, I have attempted to integrate the descriptive and theoretical approaches, by considering the diachronic development of polysynthesis in Nahuatl as a sequence of synchronic grammatical stages, parallels of each of which are still observable in other Uto-Aztecan languages.
CHAPTER 2
THE UTO-AZTECAN LANGUAGE FAMILY

2.1. The Uto-Aztecan languages and their classification

The Uto-Aztecan language family is one of the largest and most well-established language families of the Americas. In the 18th century the Uto-Aztecan languages were found over a wide geographical area from the Great Basin of the United States south to El Salvador, and from California in the west to the plains of Texas in the east. Although connections among the languages within the family have been observed since at least Buschmann (1859) (Campbell 1997), the genetic validity of the Uto-Aztecan family was not convincingly established until Sapir (1913-1919[1915])'s use of the classical comparative method to relate Southern Paiute to Nahuatl. While some prominent scholars, e.g. A. L. Kroeber, were already convinced of a relationship between the “Shoshonean” languages, Nahuatl, and other languages of Mexico, it was Sapir’s careful establishment of phonological and morphological correspondences between these languages that solidified the classification. While there is little to no current disagreement that the Uto-Aztecan family is a valid linguistic grouping, the sub-groups within Uto-Aztecan have always been uncertain and remain the subject of much current debate. In recent work both the comparative method and lexicostatistical methods have been used in attempts to work out the internal relationships among these languages.

The primary division of Uto-Aztecan is into two geographical branches: Northern Uto-Aztecan (NUA) and Southern Uto-Aztecan (SUA). The central area of contention in current debate is whether these areal splits can also be considered to be genetic
designations. A typical breakdown of the family is given in Figure 2.1 (adapted from Miller 1984), where points of dispute mentioned in this text are given with question marks.

<table>
<thead>
<tr>
<th>The Uto-Aztecan Language Family</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Northern Uto-Aztecan (??)</strong></td>
</tr>
<tr>
<td><strong>Numic</strong></td>
</tr>
<tr>
<td>a. Western Numic: Mono, Northern Paiute</td>
</tr>
<tr>
<td>b. Central Numic: Panamint, Shoshone, Comanche</td>
</tr>
<tr>
<td>c. Southern Numic: Kawaiisu, Ute (Chemehuevi, Southern Paiute, Ute)</td>
</tr>
<tr>
<td><strong>Tubatulabal</strong></td>
</tr>
<tr>
<td>Tubatulabal</td>
</tr>
<tr>
<td><strong>Takic</strong></td>
</tr>
<tr>
<td>a. Serrano-Gabrielino</td>
</tr>
<tr>
<td>(1) Serran: Serrano, Kitanemuk</td>
</tr>
<tr>
<td>(2) *Gabrielino (Gabrielino, Fernandino)</td>
</tr>
<tr>
<td>b. Cupan</td>
</tr>
<tr>
<td>(1) Cupeño, Cahuilla</td>
</tr>
<tr>
<td>(2) Luiseño</td>
</tr>
<tr>
<td><strong>Hopi</strong></td>
</tr>
<tr>
<td>Hopi</td>
</tr>
<tr>
<td><strong>Southern Uto-Aztecan (??)</strong></td>
</tr>
<tr>
<td><strong>Sonoran (??)</strong></td>
</tr>
<tr>
<td>a. Tepiman: Upper Piman (Tohono O’odham, Akimel O’odham, *Nevome), Pima Bajo, Northern Tepehuan, Southern Tepehuan</td>
</tr>
<tr>
<td>b. Taracahitan</td>
</tr>
<tr>
<td>(1) Tarahumaran: Rarámuri (Tarahumara), Guarijío</td>
</tr>
<tr>
<td>(2) Opatan: *Opata, *Eudeve</td>
</tr>
<tr>
<td>(3) Cahitan: Yaqui, Mayo</td>
</tr>
<tr>
<td>c. Tubar</td>
</tr>
<tr>
<td>Tubar</td>
</tr>
<tr>
<td><strong>Corachol-Aztecan</strong></td>
</tr>
<tr>
<td>a. Corachol: Cora, Huichol</td>
</tr>
<tr>
<td>b. Aztecan</td>
</tr>
<tr>
<td>(1) Pochutec</td>
</tr>
<tr>
<td>(2) General Aztec: Pipil, Aztec (many varieties)</td>
</tr>
</tbody>
</table>

* = an extinct language

Figure 2.1. The Uto-Aztecan languages & major sub-groups (adapted from Miller 1984)
A map showing the distribution of the Uto-Aztecan languages is given as Figure 2.2.

Heath (1978) argues for NUA as a distinct sub-group by presenting morphosyntactic evidence (in the "*na-class" verbs) of shared derived characteristics that must have developed after the break-up of Proto-Uto-Aztecan, thus implying that there was once a Proto-Northern Uto-Aztecan (PNUA). Heath presents several converging lines of

---

1 Used by permission of the Smithsonian Institution.
evidence for this classification: i-ablaut; certain thematic suffixes (*-ki-, *-tai-); a unique complex transitive verb paradigm; as well as unique participials. All of these are lacking in the southern languages and thus point to some common ancestry for the northern languages. Manaster Ramer (1992)'s sound law of lenition of medial **-c- to *-y- also uniquely identifies the NUA languages, and thus we conclude that NUA is a valid genetic group.

The comparative linguistic evidence is considerably weaker for the proposed Southern Uto-Aztecan sub-groups, precisely because the phonological evidence is not as strong for these languages. Voegelin, Voegelin, and Hale (1962) propose three sub-branches within Uto-Aztecan: “Shoshonean” (corresponding to what we now know as Northern Uto-Aztecan), Sonoran, and Aztecan. However, they were unable to provide any phonological developments that uniquely identified the Sonoran group of languages.² Using lexicostatistical evidence (i.e. relative cognate density in basic vocabulary), Miller (1984) and Cortina-Borja and Valiñas (1989) support this model for the Southern languages of the family, claiming that there is a definitive SUA branch divided into Sonoran and Aztecan.³ Because of the relative low percentage of shared cognates among the NUA languages, Miller rejected the validity of NUA as a sub-grouping. However, in light of the comparative evidence and shared developments in these languages, Miller’s lexically-based rejection must be taken less seriously than the morphological and

---

² This tripartite view had been traditionally held amongst Uto-Aztecanists, but had been rejected by Kroeber (1934) and Whorf (1935), the latter of which also rejected “Shoshonean”.
³ In an even earlier study, Hale (1964) also presented evidence for a Sonoran sub-branch, but without linking Sonoran to Aztecan in a distinct SUA sub-group.
phonological evidence that has been presented. The failure of the NUA languages to breach Miller’s 40% cognate density threshold in his 100 word sample is not conclusive evidence that these languages are not the descendents of a shared common ancestor.

At our present state of knowledge, no strictly phonological or morphological evidence supports the existence of SUA as a distinct branch from PUA. Silver and Miller (1997) provide three sound laws which they claim constitutes such evidence. Two of the sound correspondences (NUA /ŋ/ ~ SUA /n/ and NUA /n/ ~ SUA /l, r/) are equally likely (if not more so) to have been innovations in NUA, thus Silver and Miller’s proposal that they are unique to SUA is not definitive. Their third proposed sound law, PUA **i → PSUA *e, is not true for an entire branch of SUA—the Tepiman languages retain the /i/ vowel, therefore this is not a candidate for a legitimate SUA sound law. Although there is no evidence for grouping the Southern Uto-Aztecan languages into a unit that can be supposed to exist later than the break-up of Proto-Uto-Aztecan itself, we can refer to “SUA” as a geographical designation to distinguish those languages which are not a part of NUA.

Within SUA, although the lexicostatistical evidence discussed by Miller (1984) and Cortina-Borja and Valiñas (1989) leads to grouping Cora and Huichol (“Corachol”) with “Sonoran”, Campbell and Langacker (1978) give some suggestive phonological evidence that would link this sub-grouping to Aztecan. This latter position is reflected in Figure 2.1. Also, following Miller (1984)’s lexicostatistical results I have included the extinct languages Opata and Eudeve in the Taracahitic sub-group. These extinct languages are known to us only through colonial-era manuscripts recorded by various Spanish
missionaries who were in contact with these groups during and after the Conquest. Shaul (2001) has suggested that while the lexical make-up of Eudeve is Taracahitic, its grammar is more reminiscent of a Tepiman language, and he proposes that Eudeve might be a “creoloid”, a “(Te)piman reading of Opata” (p. 231).

A final area of current dispute is in regard to the correct location of Tubar within the family. Stubbs (2000, 2003) points out that although it has lexicostatistical affinity with the Sonoran languages, Tubar shares phonological similarities with languages in Sonoran as well as in Corachol-Aztecan, thus Tubar appears to be a hybrid language with uncertain (and possibly unascertainable) origin.

Although no diagnostic phonological evidence has so far been adduced which would lead conclusively to the questionable SUA sub-groupings presented in Figure 2.1, if the languages in these groups (i.e. “Proto-Southern-Uto-Aztecan” or “Proto-Sonoran”) do come from a single common ancestor language then they should show some shared innovations in all areas of their grammatical systems (Thomason and Kaufman 1988). The next logical step would be to identify morphological and/or morphosyntactic innovations among these languages, and it is in the possibility of discovering these that future studies in comparative Uto-Aztecan morphosyntax might eventually contribute to issues in Uto-Aztecan sub-grouping.

2.2. Uto-Aztecan Grammar

Despite their genetic relationship the Uto-Aztecan languages show a range of morphological types, from isolating to polysynthetic. The most comprehensive
discussion of the reconstruction of the morphology and syntax of PUA remains

Langacker's *Overview of Uto-Aztecan Grammar* (Langacker 1977a). Langacker gives the following reconstructions for the Proto-Uto-Aztecan phoneme inventory:

<table>
<thead>
<tr>
<th>*p</th>
<th>*t</th>
<th>*c</th>
<th>*k</th>
<th>*k'</th>
<th>*ʔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>*s</td>
<td>*n</td>
<td>*l</td>
<td>*w</td>
<td>*y</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2.1: Proto-Uto-Aztecan consonants* (Langacker 1977a: 22)

<table>
<thead>
<tr>
<th>*i</th>
<th>*i</th>
<th>*u</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>*o</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 2.2: Proto-Uto-Aztecan vowels* (Langacker 1977a: 22)

Langacker's reconstruction (and the general consensus) for the neutral word order of PUA was S(ubject)-O(bject)-V(erb), since most Uto-Aztecan languages display such OV traits as predominance of postpositions, the use of derivational suffixes, and head-final relative clauses. In addition to fulfilling the principle of economy in historical reconstruction (i.e. if most of the daughter languages are SOV, it would be simpler to reconstruct a few later changes to explain the deviations), a verb-final proto-language would also explain some of the more typologically peculiar aspects of some Uto-Aztecan languages. For instance, in the typological literature, one of the few proposals for "exceptionless universals" (i.e. such universals as are attributed to work like Greenberg 1963) is that V-initial languages have prepositions and not postpositions. This observation has led to common claims such as the following quotation from Song (2001):

"Verb-initial languages... are always found to be equipped with prepositions, not with
postpositions” (p.3). Following up on the endnote that appears amended to this claim we see Song’s slight retreat from such a strong position, wherein he states that “this example is often cited as an exceptionless universal. But it is correct to say that there are a few verb-initial languages with postpositions”, including the Uto-Aztecan languages Northern Tepehuan and Cora (p.46). If these two languages descended from an SOV language, then the “universality” of the typological statement certainly has more empirical force, since the exceptions here could be explained by subsequent developments in these particular languages. Time will tell if they develop prepositions or not. In addition, some Uto-Aztecan languages have taken on prepositions; Hill and Hill (2004) discuss the borrowing of the Spanish preposition de (and others) in Nahuatl, a result of the long period of contact between those two languages.

In the following discussion I will highlight only a few key points of comparative Uto-Aztecan grammar, those that will serve as useful background for later chapters. Our focus here will be on general sentence structure, DPs and case-marking, and subject and object clitics. Other areas of comparative morphology will be highlighted in later chapters, including general properties of word-order and case-marking, reduplication, noun incorporation and the formation of denominal verbs, and instrumental prefixes.

2.2.1. Overview of Uto-Aztecan Syntax

As mentioned above, Langacker (1977a) reconstructs SOV as the neutral word order of PUA; however, it is assumed that topicalization and other pragmatic factors could lead to variation in this basic order. In addition, Langacker reconstructs for PUA what he calls
“pronoun copies”, a category of grammatical elements that includes both bound elements (affixes or clitics) as well as resumptive and presumptive pronouns. According to Langacker, “the four basic grammatical constructions for which pronoun copies are invoked in UA are those involving possessives, postpositions, object agreement on the verb, and subject agreement on the verb or in a clitic group” (1977a: 27). These “copies” can stand alone—“when the pronominal element is itself sufficient to identify the referent, the antecedent can be omitted, in which case we do not have a pronoun ‘copy’ in the strictest sense of the term” (p.27). We will return to the issue of the grammatical marking of subjects and objects below. In the meantime, illustration of the use of pronominal elements with possessives and postpositionals is given below in (1) and (2), respectively:

(1) **Mono**

\[
\text{pahapi"ci-na nii a-"ki"ki-na a-na-"pu"ni-"ti}
\]

bear-ACC I its-feet-ACC it-nearly-see-TNS

‘I saw the bear tracks’

(Langacker 1977a: 25 [7])

(2) **Southern Paiute**

\[
\text{a-tu"k"a kani-a}
\]

it-under house-ACC

‘under the house’

(Langacker 1977a: 27 [15])

Langacker reconstructs the accusative case-marking on the antecedent of such constructions, seen above, but this has not been retained in all of the daughter languages. Langacker also reconstructs *inversion*, where the copy precedes its antecedent, as in (2), as well as *discontinuity* in such constructions, where a constituent is interrupted by some other grammatical material, as in (1).
In addition to these pronominal “copies”, Langacker reconstructs three “unspecified argument” prefixes for PUA. These were *ta- ‘unspecified subject’; *ti- ‘unspecified object’; and *ni-, “an unspecified human subject presupposed to be coreferential to the object (also unspecified)” (p.46). According to Langacker, 

Most daughters have lost one or more of these prefixes, and when retained they have often been modified. Only Shoshoni and Classical Aztec retain all three; *ni- has become passive and reflexive in Shoshoni, and in Nahuatl *ti- and *ta- (reflected regularly as tee-* and tla-) have become realigned in function and mark unspecified human and non-human objects respectively through a change dating back to Proto Aztecan. (p.46)

PUA also had reflexive, reciprocal, passive, and impersonal morphological marking. Langacker (1976) thoroughly traces the historical development of these morphemes throughout the family.

In terms of overall verb structure, as Langacker (1977a: 157-9) points out, there is great variation in the complexity of the verbs across the Uto-Aztecan languages. Serrano, for instance, has “relatively little productive verb morphology and comparatively few verbal affixes, and long strings of affixes are simply not found” (Langacker 1977a: 157). Huichol, on the other hand, represents the other extreme. Grimes (1964) “sets up fifteen prefix position classes and five suffix position classes in a morpheme-order chart”, and “as a group the UA languages tend to be more like Huichol than like Serrano in regard to verbal complexity”, although Huichol is unusual for having such an abundance of prefixes (Langacker 1977a: 157).
Langacker concludes that the typical UA verb structure is “largely agglutinative, with boundaries between affixes normally easy to discern, comparatively little fusion of adjacent affixes, and little in the way of portmanteau morphemes” (p.158). With respect to the order of affixes, Langacker notes that “there is a tendency... for the more noun-like affixes (e.g. pronominal or instrumental prefixes) to precede the stem and for more verb-like affixes to follow the stem; this is not unrelated to the verb-final character of the family” (p.158).

Finally, Uto-Aztecan is well-known for its verb stems that are suppletive for number, wherein “suppletion is consistently with respect to the number of the subject with intransitive verbs and the number of the object with transitive verbs”, traits that “can definitely be reconstructed for [PUA]” (p.127).

We turn now to a brief discussion of subject and object DPs across the UA family.

2.2.2. DPs and Case-marking

It is agreed that PUA had a nominative/accusative case-marking system, which has been retained in most of the daughter languages. Langacker (1977a) reconstructs *-a as the marker of accusative, while nominative was unmarked (pp. 82-3).

In addition, PUA had what has been referred to as the “absolutive” case morpheme, although this nominal marking has nothing to do with the ergative/absolutive case-marking familiar from other languages. According to Langacker, the Uto-Aztecan absolutive “is an ending with no apparent semantic value that appears on nouns in citation forms but may drop when a noun is subjected to various morphological
processes, such as affixation, compounding, or reduplication” (p.77). More neutrally, we may refer to this morpheme as a marker of non-possessed nominals (NPN’s), as in J. Hill (in press).

Langacker reconstructs *-ti as the PUA absolutive, which is thought to have been retained before the accusative *-a and possibly also the plural *-mi (1977: 77). The combination of absolutive *-ti plus accusative *-a has been reanalyzed as a single “non-nominative” case marker in some languages (e.g. Yaqui –ta ‘non-nominative, non-plural’; Hopi –ta ‘ACC’). The absolutive has more or less retained its NPN-marking function in at least Tübatulabal, Classical Aztec, and Serrano, which show the sets of suffixes {–l, –t}, {–tl, –tli, and –l(i)}, and {–t, –c, and –c}, respectively, all of which can be traced straightforwardly back to PUA *-ti. Examples of the UA absolutive from each of these languages are given in (3), (4), and (5) below, which show complementary distribution between the NPN marker and possessive pronoun (a vs. b), and in Classical Aztec and Serrano also between the NPN marker and postpositions (4b’ and 5c):

(3) Tübatulabal
   a. hanii-l
      house-NPN
      ‘house’
   a’. tabaaya-l
      chipmunk-NPN
      ‘chipmunk’
   b. hanii-n
      house-his
      ‘his house’

   (Steele 1979: 445 [46])

   (Steele 1979: 445 [6])

---

4 According to Jane Hill (personal communication), the PUA absolutive is retained in Takic more broadly.
5 According to Steele (1979), the absolutive in Tübatulabal does not alternate with postpositions because “there are no postpositional suffixes, only independent adpositional elements” (p. 465).
(4) Classical Aztec
a. *tilmaʔ-tli*
   blanket-NPN
   ‘blanket’
   (Steele 1979: 445 [5])

b. *no-tilmaʔ*
   my-blanket
   ‘my blanket’
   (Steele 1979: 446 [7, 8])

(5) Serrano
a. *kii-č*
   house-NPN
   ‘house’
   (Steele 1979: 465 [29])

b. *mi-ki*
   your-house
   ‘your house’
   (Steele 1979: 466 [31])

c. *kii-kaʔ*
   house-to
   ‘to the house’
   (Steele 1979: 466 [33])

The Uto-Aztecan languages vary with respect to how this “absolutive” has developed. In addition to the reanalysis of the Uto-Aztecan absolutive as a non-nominative case-marker in Yaqui and Hopi, Steele (1979) shows that while Luiseño has absolutive suffixes, their forms are more complex than can be attributed to the PUA absolutive, and she posits more than one underlying morpheme serving this function in that language.

Southern Paiute has the suffixes *-pi* and *-pē*, the former of which is used for “referring chiefly to animals, topographical features, and objects” (Sapir 1930: 113), and the latter “on nouns which are the names of plants” (Steele 1979: 468). Although Steele cautions that “although these suffixes may disappear when the noun is possessed, there does not appear to be the regularity to this alternation that is attested, e.g. in Luiseño” (p. 468), she suggests that these suffixes function analogously to the PUA absolutive. Finally, some
languages have lost “both the form of the old absolutive and anything which resembles its patterning” (p. 469), with only remnants of the form appearing sporadically in Tohono O’odham, e.g. “between certain postpositions and the noun to which they attach is a –t or –č, remnants of the old absolutive” (p. 470).

Uto-Aztecan adjectives and determiners, etc., present an interesting domain for future inquiry into the domain of DPs. According to Langacker (1977a), “it is not uncommon for most adjectives in a UA language to be derived (synchronically or diachronically) from members of other classes and to have recognizable derivational markings, leaving only a few apparently underived adjectives” (p. 66). These underived adjectives are often difficult to distinguish from stative verbs, as with the following examples from Hopi:

(6) Hopi
   a. moosa qati (Langacker 1977a: 66 [1])
      cat sit
      ‘The cat is sitting’
   b. moosa qōōca (Langacker 1977a: 66 [2])
      cat white
      ‘The cat is white’

Such underived adjectives can be used attributively with appropriate morphology, as with relativization or nominalization (e.g. 7a vs. 7b):

(7) Hopi
   a. moosa qati-wta-qa (Langacker 1977a: 66 [3])
      cat sit-STAT-AG
      ‘sitting cat’
   b. qōōca-mosa (Langacker 1977a: 66 [4])
      white-cat
      ‘white cat’
A full comparison of adjectives and other nominal modifiers across the family goes beyond the scope of the present work; see Langacker (1977a: 66-69) for some descriptive generalizations.

Adjectives in some UA languages, however, can be inflected for accusative case (e.g. Hopi and Yaqui). Based on a comparative analysis of Hopi with other Uto-Aztecan languages, K. Hill (2001) reconstructs the following number and case morphemes for PUA:

(8) Case and Number Morphemes in Proto-Uto-Aztecan (K. Hill 2001:306)

absolutive accusative non-pluralized: **-ta
plural (two distinct suffixes): **-ti, **-mi
accusative
adjectival nominative
adjectival accusative

2.2.3. Subject and Object Clitics

The Uto-Aztecan languages typically allow for (if they do not require) the marking of subject and/or object relations by means of clitics or affixes. Steele (1977) examines the status of bound pronominal elements across Uto-Aztecan and makes two distinct arguments regarding their development from PUA. First, these pronominal elements are historically derived from free pronouns, and I will give some of her evidence in support of this view below. The second claim that Steele makes is more controversial. She summarizes her position as follows: “all the pronouns which are proclitic (or prefixed) to the verb are derivative from other bound pronominal forms, most from previously cliticized second position pronouns” (p.539), of the type currently seen in languages like
Tohono O’odham. In contrast, I will be arguing that there is little reason to believe this second claim.

First, to address the issue of the etymology of clitic and affixal pronominal elements, Steele gives a catalogue of the subject clitic/affixal elements in comparison to the free subject pronouns of the various languages, which are transparently cognate across the family. I gave a representative sampling from across the family below:

(9) **Tübatulabal**

<table>
<thead>
<tr>
<th>Clitic Pronouns</th>
<th>Independent Pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sg.</strong></td>
<td><strong>pl.</strong></td>
</tr>
<tr>
<td>1 g(i) gila</td>
<td>(gi)luuc</td>
</tr>
<tr>
<td>2 b(i) bu</td>
<td>m(u)</td>
</tr>
<tr>
<td>3 Ø da</td>
<td></td>
</tr>
</tbody>
</table>

(10) **Serrano**

<table>
<thead>
<tr>
<th>Clitic Pronouns</th>
<th>Independent Pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sg.</strong></td>
<td><strong>pl.</strong></td>
</tr>
<tr>
<td>1 ni čimi</td>
<td></td>
</tr>
<tr>
<td>2 či čimi</td>
<td></td>
</tr>
<tr>
<td>3 vi mi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(11) **Tohono O’odham**

<table>
<thead>
<tr>
<th>Clitic Pronouns</th>
<th>Independent Pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sg.</strong></td>
<td><strong>pl.</strong></td>
</tr>
<tr>
<td>1 ñ č</td>
<td></td>
</tr>
<tr>
<td>2 p m</td>
<td></td>
</tr>
<tr>
<td>3 Ø Ø</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6 In Tübatulabal, “independent subject pronouns are formed from the clitic pronouns with the prefixation of -in; the first singular form is, however, irregular” (Steele 1977: 567).
(12) Yaqui

Clitic Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ne</td>
<td>te, itom</td>
</tr>
<tr>
<td>2</td>
<td>e, en</td>
<td>em, ?eme</td>
</tr>
<tr>
<td>3</td>
<td>ø</td>
<td>ø</td>
</tr>
</tbody>
</table>

Independent Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>nelepo</td>
<td>?itepo</td>
</tr>
<tr>
<td>2</td>
<td>?empo</td>
<td>?eme?e</td>
</tr>
<tr>
<td>3</td>
<td>?aapo</td>
<td>bempo</td>
</tr>
</tbody>
</table>

(13) Huichol

Clitic Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ne</td>
<td>te</td>
</tr>
<tr>
<td>2</td>
<td>pe</td>
<td>ze</td>
</tr>
<tr>
<td>3</td>
<td>ø</td>
<td>we/me</td>
</tr>
</tbody>
</table>

Independent Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ne</td>
<td>taame</td>
</tr>
<tr>
<td>2</td>
<td>?eekii</td>
<td>zeeme</td>
</tr>
<tr>
<td>3</td>
<td>?iiki</td>
<td>?iime (near)</td>
</tr>
<tr>
<td></td>
<td>miiki</td>
<td>miimi (distant)</td>
</tr>
<tr>
<td></td>
<td>?iya</td>
<td>(general)</td>
</tr>
</tbody>
</table>

(14) Classical Aztec

Clitic Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ni</td>
<td>ti</td>
</tr>
<tr>
<td>2</td>
<td>ti</td>
<td>an</td>
</tr>
<tr>
<td>3</td>
<td>ø</td>
<td>ø</td>
</tr>
</tbody>
</table>

Independent Pronouns

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ne?waatl</td>
<td>te?waan(tin)</td>
</tr>
<tr>
<td>2</td>
<td>te?waatl</td>
<td>ame?waan(tin)</td>
</tr>
<tr>
<td>3</td>
<td>ye?waatl</td>
<td>ye?waan(tin)</td>
</tr>
</tbody>
</table>

Most UA languages have subject pronoun clitics; the languages that do not, according to Steele, are Northern Paiute, Shoshone, and Hopi. Steele regards the pronouns of Mono as “somewhere between clitics and independent forms. There is one set of pronouns in the language; these pronouns generally occur in sentential second position” (p. 541), as in the following example:

(15) Mono

nopihweeh nii miyawaih
to:home I will:go
‘I shall go home’
However, certain grammatical elements, like the modal *hsahqwa* and the conjunction *po' o*, are “more necessarily sentential second position elements than are pronouns” (p. 542).

These clitic subject pronouns are either second position elements or they occur as prefixes to the verb. Some languages in fact have both; Steele’s classification of the languages into the different types is given below.

<table>
<thead>
<tr>
<th>Second Position:</th>
<th>Comanche, S. Paiute, Chemehuevi, Tübatulabal, Serrano, Luiseño, Tohono O'odham</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proclitic/Prefix:</td>
<td>Kitanemuk, Cahuilla, Huichol, Classical Aztec, Pochutla</td>
</tr>
<tr>
<td>Both:</td>
<td>Cupéno, Tepecano, Tarahumara, Yaqui, Cora</td>
</tr>
</tbody>
</table>

*Table 2.3. Subject clitics across Uto-Aztecan* (Steele 1977: 542-3 [III])

According to Steele, clitic pronouns in Tepecano “can fill both positions simultaneously”, but “for the other languages with both types, the two positions are independent possibilities” (p. 543). These distinct positions can be filled concurrently; Steele gives examples of the co-occurrence of clitics in both positions in Tarahumara, Yaqui, and Cora. I include her Yaqui example below:

(16) **Yaqui**  

\[
\text{kwárentpeso dyaryota}=\text{ne ne}={}\text{kóba iani inine}
\]

‘Now I make forty pesos a day here’

The claim of Steele (1977)’s that I wish to dispute is that regarding the directionality of the development of these clitics. In a nutshell, the path of grammaticalization that
Steele posits for these elements is the attraction of free pronominals to second position, and then subsequent cliticization, *before* the clitics can attach to the verb. Thus, the languages with the second position clitics are in what we may regard as a "natural" state, whereas the languages with proclitic/prefixal elements have fully grammaticized their verbal prefixes from this original state. The mixed languages that have both are in an intermediate stage vacillating between the two other extremes.

The evidence that Steele posits to support this view include the data from Mono, which suggests to Steele that "the beginnings of the cliticization of independent pronominal forms can be witnessed synchronically..." in that language (p. 548). Thus, "it appears that the beginning of the cliticization of independent pronouns is indicated by their appearance in sentential second position" (p. 548). However, it is not clear that the Mono data are in fact evidence of the recent development of second position elements; if anything, the principle of parsimony would seem to suggest the opposite case, that Mono has more recently *lost* the clitic status of its pronominals, perhaps developing full pronouns from earlier second position clitics.\(^7\) The majority of Numic and other NUA languages have second position clitics, and, since as it was argued above these languages are more closely related to each other (via PNUA) than the SUA languages, it is not inconceivable that this property that they share could be attributed to PNUA itself. Thus, Steele's argument that Mono is only now developing a second position clitic from a full

---

\(^7\) The expected path of grammaticalization, of course, is the other way, with free pronouns developing into clitics. However, see Campbell (2001b) for a recent series of papers giving counterexamples to the "unidirectionality" thesis propounded in grammaticalization theory.
pronoun cannot be used as compelling evidence that all of the other Uto-Aztecan languages did as well.

Steele subdivides her proposed courses of development of prefixal elements from second position elements—one path is typical of Yaqui and Tarahumara, and the other is seen in the other languages. For the case of Yaqui and Tarahumara, where the second position and prefixal/proclitic elements are homophonous, Steele proposes that one is derivative from the other. She cites Mason (1923), quoting from an 18th Century grammar of Yaqui: “Velasco states as one of his most infallible rules that the pronominal subject [read clitic pronoun—SS] must be the second word or element in the sentence...” (cited in Steele 1977: 553, emphasis in original). Thus, Steele concludes that the verbal proclitic element has developed since the 18th Century. Since there is no such evidence from Tarahumara there is no clear direction of grammaticalization, but Tarahumara, unlike Yaqui, cannot have a proclitic element alone in a sentence—proclitics must co-occur with second position elements (Steele 1977: 554).

For the languages with proclitic or prefixed pronouns, “a second position clitic pronoun which is also contiguous to the verb has been reanalyzed as a proclitic to the verb” (p.554). This account presupposes that these languages must have lost their unmarked SOV word order, in order to allow for contiguity of the verb with the second position. As Steele puts it,

with a second position clitic pronoun and a (potentially) clause final verb, the clitic pronoun can be separated from the verb by a noun.

\[ S = CP \ O \ V \]
But if the object nominal is moved out from its position between the clitic pronoun and the verb, the two will be contiguous. The clitic pronoun is potentially reanalyzable as proclitic to the verb.

\[ S=CP \ V \ O > S \ CP=V \ O \quad (\text{Steele 1977: 556}) \]

Steele (1976) shows good evidence that Classical Nahuatl was in transition to a verb-initial language, finding that the verb rarely followed the object in sentences with both a subject and object, and although SVO order was relatively frequent, VSO was more so. Steele posits that the trigger for this change was the “attraction to the verb of sentence-initial modal particles” (p. 44). In this view, modal elements have intrinsic attraction to two conflicting locations within a clause—to the verb itself or to sentence-initial position; on this point see also Steele (1975). If the modals shift from verbal position, they can, in a sense, eventually “drag” the verb with them. Steele’s claim is that the other languages in Uto-Aztecan which have prefixal clitics are also lacking in rigid SOV order, so it is possible that they have undergone word order change, and reanalysis of second position AUX to verbal proclitic, similar to what is purported to have been shown with Classical Aztec.

However, Steele’s explanation of this word-order change does not take into account the sociolinguistic situation of the change in word order that Classical Aztec (and other extremely southern SUA languages) had developed from its earlier V-final state. As Campbell, Kaufman, and Smith-Stark (1986) show, V-initiality is an areal feature of Mesoamerica, and many languages of the area have taken on this property even though they may be descendents of V-final and other language types. Thus, Steele’s argument
for a causal relationship between second position and its supposed attraction for the verb is not as convincing when the larger cross-linguistic relationships among these languages are taken into consideration. Furthermore, as mentioned above, "pronoun copies" are not limited to second position clitics or verbal affixes. They also occur in possessive and postpositional constructions. Thus, there could have been a wider range of influence on the grammaticalization of these elements into fixed positions than just what is attracted to second position. In addition, this account would seem to predict that there could also be SOV languages in which the subject clitic was reanalyzed as a prefix on the object, but it is not clear that this is at all attested in Uto-Aztecan.

Finally, Steele’s hypothesis brings about an independent problematic issue: the fact that Steele presupposes that there was already a second position AUX element in PUA (see especially Steele 1979 for arguments for this reconstruction). However, only two languages exhibit “the position, composition, and relative order hypothesized to the proto-language” (Steele 1979: 474)—Luiseño and Cupeño, two very closely related languages in the Cupan group of the Takic sub-branch of NUA. A detailed critique of this proposal would be interesting but goes beyond the scope of our discussion here.

Rather than seek directionality of development between second position clitic elements and affixal agreement markers, I hypothesize instead that PUA had both possibilities for subject-marking. The distribution within the family of these disparate kinds of marking is consistent with such a proposal. Languages with both kinds are attested in both NUA (Cupeño) and SUA (Tepecano, Tarahumara, Yaqui, Cora), as are languages with only proclitic/prefixal subject marking (Kitanemuk and Cahuilla in NUA,
Huichol, Pochutla and Classical Aztec in SUA). Languages with only second position clitics are attested in both NUA and SUA, but there remains a lingering possibility that these second position clitics are an areal phenomenon, since all of those languages that have this feature are contiguous (J. Hill in press). If it is true that PUA had both kinds of marking, then it is simpler to say that the languages that retain both are unmarked, whereas the languages that have a more fixed position have “run with” one or the other options available to them, possibly at different times and under different circumstances (i.e. under this account, no unified explanation would need to be called for).

Thus far in this discussion I have focused on Steele’s account of the development of subject pronominal clitics and affixes of the two kinds that we see in Uto-Aztecan: those that occur in second position and those that are bound to the verb. Other than in terms of their probable etymology from free pronouns we have not considered object clitics at all. The topic of object “pronominal agreement” will be taken up in detail in our discussion of noun incorporation in Section III, as well as in Chapter 9, where I develop a theory of the development of obligatory polysynthesis (i.e. subject and object agreement on the verb) in Nahuatl.

2.3. Uto-Aztecan historical linguistics and cultural prehistory

One of the important contributions that historical linguistics can make to our general knowledge goes beyond the realm of linguistics proper and into more anthropological concerns regarding prehistoric contact(s) and the reconstruction of culture(s). In the long history of Uto-Aztecan studies the reconstruction of Uto-Aztecan prehistory has been a
primary concern, although there is still debate about the origins of the PUA community and the direction of its dispersal. Since the Uto-Aztecan language family covers such a wide geographical area, and since Uto-Aztecan-speaking peoples can be found in several important, distinct cultural areas, the unraveling of Uto-Aztecan prehistory in some conclusive way will be a major contribution to the study of the history of the Americas more broadly. This section will briefly outline the present state of our knowledge, focusing on the role of the historical linguistic evidence in the reconstruction of culture and the movement of populations.

Considering the geographical range of the Uto-Aztecan family at the present day, proposals for the spread of the family can be generally classified into three distinct hypotheses, which happen to more or less reflect the chronology of their proposal. We can consider these to be Northern, Central, and Southern accounts of the PUA homeland and the origins of the Uto-Aztecan languages.

The extreme Northern view was held by Zingg (1939), who proposed that the Uto-Aztecanans crossed into North America via the Bering Strait, progressively moving southward into the Great Basin, where they adopted the Basket-Maker culture. This they subsequently carried with them into southern California and as far south as Chihuahua and Sonora. The Nahua-speaking Aztecs (Mexico) were late arrivals into the Valley of Mexico. However, Zingg’s study was based on material culture alone, with minimal attention paid to the results of comparative linguistics, unlike later studies which presented more solid evidence for the origins of the Uto-Aztecan family. Hopkins (1965) presented a similar northern view, based on the assumed relationship of Uto-Aztecan
with other northern groups such as Penutian and Sahaptian. Goss (1968) dismisses Hopkins’ linguistic evidence, and in any event such remote relations for Uto-Aztecan are not currently generally accepted based on the evidence that has hitherto been presented (Campbell 1997).

A more centrally-located homeland for PUA has been the standard assumption for many years. Goss (1968) surveys several proposals for UA origins that were under consideration at that time, and correlating linguistic and archaeological results he concluded that “the linguistic evidence seems to indicate that the ‘center of gravity’ for the Utaztekan stock would lie somewhere around the Arizona-Sonora border (historic Pimic country)” (p. 17). This general model is supported by Fowler (1983), who uses biogeographical terms common throughout Uto-Aztecan, e.g. such PUA terms as **amol ‘agave sp.’ and **k’iyo ‘turkey’, to locate the PUA homeland in Northwestern Mexico and the US Southwest, possibly extending into southeastern California. This proposal indicates a large range, reflecting the possibility that PUA was composed of a dialect continuum (J. Hill 2001), although Fowler proposed that there was a distinct NUA and SUA split (1983: 246).

J. Hill (2001) has recently proposed a more radical view of PUA origins, arguing that PUA must have originated in Mesoamerica, thus proposing a Southern origins model. She argues that shared vocabulary in the maize complex indicates that PUA was spoken by an agricultural people, placing them farther south in Mesoamerica, where corn agriculture developed, than has been previously believed. Under Hill’s account, the northward migration of Uto-Aztecan peoples was induced by population pressures
concomitant with the spread of cultivation. Although the northern UA groups do not have agriculture, Hill argues that they have abandoned their agricultural tradition, yet retain some terminology associated with the maize complex. Crucially, Hill points out that this origin of PUA is not inconsistent with the biogeographical reconstruction presented by Fowler (1983). Dakin (2003, 2004) and Dakin and Wichmann (2000) have recently supported Hill's model by arguing that certain old Mesoamerican borrowings have Uto- Aztecan origins, rather than vice versa, suggesting that Uto-Aztecan languages have been in Mesoamerica longer than is allowed for by the earlier models of dispersal.

The Northward spread of the Uto-Aztecan languages, from either Central or Southern origins, is consistent with Dixon (1997)'s recent discussion of punctuated equilibrium in historical linguistic change. Under Dixon's account, the family tree model of linguistic relationships is only of use in sporadic, discrete splits, which come about in historical circumstances of abrupt change, such as with rapid population dispersal. This seems to be exactly what we see in NUA, where there is general agreement about the intermediate familial relationships. In SUA, on the other hand, presumably where there has been more intermixture of languages for longer periods of time, we see more of a mesh than a tree (Miller 1984). This is exemplified by claims of certain languages being creoles (e.g. Eudeve) (Shaul 2001), as well as other languages which seem to be intermixed (e.g. Tubar) (Stubbs 2000), in SUA.

Further inquiry into Uto-Aztecan prehistory will no doubt shed further light on prehistoric linguistic relationships. Following the approach of Thomason and Kaufman (1988), genetic relationships should be evident not only in shared vocabulary and
phonological retention and innovation, but also in other areas of grammar as well. This would apply not only to Uto-Aztecan-internal relationships, but to larger-scale relationships between Uto-Aztecan and other linguistic groups. However, there remains an open question as to how far the comparative method can take us in establishing remote relations. See Campbell (1997) for a skeptical review of proposals that attempt to link Uto-Aztecan to other stocks. In Chapter 3 we will discuss in further detail the use of the comparative method in historical syntactic reconstruction.

2.4. Uto-Aztecan and linguistic theory

This dissertation was born of a particular attitude toward the scientific study of language, a position that holds that linguistic theory must be held accountable to the diversity of the world’s languages. In this view, theoretical hypotheses about the nature of language, whether synchronic or diachronic, must be tested against a wide range of languages and language types. This position entails what Ken Hale has called the “the confirmatory function of linguistic diversity” (Hale 2000: 168).

The flip side of this coin is that it should also be the case that insights drawn from linguistic theory can have some effect on the way that we view specific languages, especially when seeking to explain the limits of variation that we observe across languages. This is particularly crucial for languages that are very closely related, but demonstrably different. Two critical questions in such cases are: what did these languages originally share, and how did these languages become different?
This chapter has presented some of the basic findings of comparative Uto-Aztecan morphology and syntax that form the backdrop on which the major sections of this dissertation are founded. Much of this work was done from the typology-based perspective of comparative linguistics that was en vogue in the 1970s, including especially the work of Langacker (1976, 1977a) and Steele (1975, 1976, 1977, 1979). Although the specific questions that are asked below are framed within a particular theoretical backdrop (which will be outlined in detail in Chapter 3), the results of this investigation also have empirical ramifications that go beyond the narrow scope of theory-central questions. I hope that future study in comparative Uto-Aztecan morphosyntax can take a larger role in questions of theory-testing and theory-building, in addition to illustrating the ways in which the theoretical perspectives adopted below lead to analyses that tell us something about historical change in Uto-Aztecan.

From the perspective of the history of linguistic study, it is ironic that Uto-Aztecan has not played a larger role in comparative syntax from the generative perspective. As Steele summarized her overview of Uto-Aztecan historical linguistics in the seminal volume edited by Campbell and Mithun in 1979, she concluded as follows:

Future research in Uto-Aztecan assuredly will attempt to fill the holes in our knowledge of the language family; future research in Uto-Aztecan will directly bear on theoretical issues only to the extent that Uto-Aztecanists make that aspect of their work clear. As an assessment of the state-of-the-art in Uto-Aztecan, then, this paper argues, by example, for an attitude towards research, an attitude that looks beyond the specifics of the field. The argument is two-edged. While Uto-
Aztecanists must be concerned with general theoretical questions, those who consider themselves hard-core linguistic theorists must take careful account of research outside the usually very small data base that informs their theories, research like that already being done in Uto-Aztecan. (p. 507)

The irony is that this statement was made just before the dawn of the Principles and Parameters era, typically attributed to Chomsky's (1981) *Lectures on Government and Binding*, where the central goal of theoretical inquiry in syntax was shifted to the explanation of cross-linguistic variation, based upon a limited set of variables which were intended to be discerned through cross-linguistic investigation. However, although data from Uto-Aztecan languages has been used in linguistic theorizing from this perspective (e.g. the discussion of Tohono O'odham as a non-configurational language in Jelinek 1984), data specifically from comparative Uto-Aztecan has played little role in the development of this paradigm.

The central goal of this dissertation is to apply the results of comparative inquiry in Uto-Aztecan to central questions in current generative linguistic theory, as presaged by Steele (1979). We turn now to the theoretical background and assumptions which underlie the major empirical and theoretical substance of later chapters.

---

8 A similar focus in generative phonology was not developed until the advent of Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993).
3.1. Introduction

The theoretical aspects of this dissertation are based within the framework of Distributed Morphology (DM), and one of the points that I explore in later chapters is the application of this model to issues in problems of comparative and historical morphosyntax in Uto-Aztecan. In this chapter I give background on this model and its central tenets in section 3.3. In section 3.2, however, I will first discuss some more general assumptions behind my approach to the study of language, in order to justify the use of a generative theoretical framework in historical linguistic investigation. Section 3.4 will consider previous approaches to diachrony in syntax and will discuss the issue of syntactic reconstruction.

3.2. General assumptions about the nature of language and how we should study it

Following the general assumptions of the generative linguistic tradition, dating back to at least Chomsky (1965), I assume that at least one aspect of the study of language involves the study of the mind—those aspects of linguistic knowledge that generally lie beyond our conscious awareness but which nevertheless involve some aspect of what we know. Part of the study of “what we know when we know a language” involves the formation of grammaticality judgments. I assume that all speakers have such intuitions about their native language(s), and information about the structure of any language can be sought...
through the intuitions of native speakers. This is not to say, however, that there are not problematic issues with the employment of this methodology. Schütze (1996) discusses many important aspects of the scientific grounding of the procedures used in eliciting grammaticality judgments. In the end, however, I agree with Bever and Langendoen (1971)’s position that forming grammaticality judgments is just one of several aspects of linguistic behavior. As they phrase it, “It is clearly the case that the activities of talking and listening can obscure much of a person’s linguistic knowledge; but judgments about potential sentences are also behavioral manifestations of linguistic knowledge, and as such are not different in principle from the more direct uses of linguistic structure” (p. 433).

Following the distinction between I(nternal)-language and E(xternal)-language (Chomsky 1986), or even langue and parole (Saussure 1916), the object of study of relevance here is that of linguistic competence, which is the mental reality that underlies linguistic performance (Chomsky 1965). This study of language reveals significant aspects of the mind, and the study of linguistic variation leads to the potential of ascertaining important constraints and limitations on possible mental states. The assumption here is that syntactic variation is finite and delimitable. In my view, there is complementarity between studies of competence and studies of performance. Thus, I see no essential conflict between seeking native speaker judgments about particular syntactic constructions of a given language and more usage-based corpus studies, as performed, for example, in the classical Boasian tradition of collecting texts to make inferences regarding grammatical structure. Whereas the latter kind of study tells us important
things about the use of language in particular discourse contexts, the former type of study can lead to understanding other aspects of possible utterances. To quote Newmeyer (2003)'s recent dictum, "grammar is grammar and usage is usage."

Since any child can acquire any language if given the relevant exposure, I follow the general approach to linguistic theorizing that seeks to ascertain the limits of possible variation among languages, under the assumption that children begin from the same starting point. This approach assumes that all languages are fundamentally of the same nature, and that the radical surface variation that is so obvious to even casual cross-linguistic investigation is the result of a finite number of permutations of a more general underlying structure, or "Universal Grammar". This view contrasts maximally with the recent development of an approach to language proffered by Croft (2001), "Radical Construction Grammar", which holds that "there is no universal syntactic template to which the grammars of all languages conform. Nor is there an inventory of universal syntactic categories, relations or even constructions which the grammars of all particular languages draw from" (Croft 2001: 61).

Although this dissertation is couched in terms of recent generative theorizing, three recent proposals within this framework are not followed here, nor are they necessary for a universalist approach to underlying linguistic structure.

First, I do not assume the inverse of Radical Construction Grammar, i.e. a "Radical UG", that would posit a universal linguistic architecture for all languages. For example, I do not assume that all functional projections necessarily appear in the same order in all languages, as recently posited for adverbs by Cinque (1999). However, I do agree with
Cinque that it may be desirable to make this assumption at the outset as the null hypothesis, if only to refute it later through empirical investigation. I assume that the range of variation is limited and, crucially, learnable.

Second, I do not assume identical universal branching of syntactic structure; i.e. I do not follow Kayne (1994)'s Linear Correspondence Axiom (LCA), requiring that all branching be rightward. Thus, in my discussion below of the head-final Uto-Aztecan languages, I assume some version of the Head Directionality Parameter (Baker 2001: 68), and I assume leftward branching for these languages. There may be theoretical reasons to force rightward branching, but these go beyond the scope of our consideration here (see Julien 2002 for a recent treatment of head-finality under LCA assumptions). However, I do adopt the (standard) position that all branching is binary.

Finally, although I am working within a related framework, I do not follow the central notion of the recent “Minimalist Program” (Chomsky 1995) that language is a “perfect” system; this issue, and whether it is desirable or not, plays no role in the discussion to follow. However, I do agree with the general Occam’s razor point that the best theory of language is the one that posits the fewest mechanisms to derive the greatest empirical coverage. In the course of seeking a theory of this nature, it is also important that the theory be falsifiable. We turn now to a discussion of the tenets of the theory in which the ensuing analyses will be framed.
3.3. The Pervasive Syntax Perspective and Distributed Morphology

The framework of Distributed Morphology (DM) was established by the pioneering work of Halle and Marantz (1993, 1994), which has subsequently inspired a wide range of research in theoretical morphology. In this section I will briefly outline the major tenets of the approach and address why I think it is a fruitful framework for research into comparative morphosyntax. This discussion follows closely the overview of the framework provided by Harley and Noyer (1999). First, however, we must contrast DM with other current approaches to the interface of syntax and morphology, which can be differentiated from DM on two crucial grounds—underlying assumptions about the nature of morphology (whether it is piece-based or process-based) and the general attitude towards the Lexicalist Hypothesis of Chomsky (1970).

Linguistic theories of morphology can be grouped into two broad categories which are defined according to one’s view of the nature of morphology. These contrasting approaches to morphology are known as item and arrangement (IA) vs. item and process (IP) approaches (Hockett 1958), or, in other terms, piece-based vs. process-based approaches. The former conceives of morphemes as “things”, or “pieces”, “and morphology is simply the concatenation of things, so is viewed as formally agglutinative” (Spencer 1998: 123). The latter view “holds that morphology should be regarded as a set of processes acting on stems or words to produce new stems, words or word forms” (Spencer 1998: 124). One of the more prominent recent proponents of an IP model of morphology is Anderson (1992). A third view, the Word-and-Paradigm approach (WP), is also non-agglutinative, and it proposes morphosyntactic paradigms as the central
notion in the grammar. In this approach, "each inflected form has (at least) one morphosyntactic description (for example 'past tense form' or 'dative singular of the masculine/neuter adjectival form') and the grammar then makes available paradigms that specify the formatives which correspond to these categories" (Spencer 1991: 52).

One criticism of IP and WP approaches is their potential lack of a principled limitation on processes and members of paradigms. Under these approaches it is possible that recurring morphological pieces, e.g. English plural -s, could be treated as only coincidentally (or accidentally) related, and the theory does not disallow radical phonological processes, even though they do not occur in natural languages (cf. Marantz 1982's examples of unattested reduplication patterns of the form $C_1V_2C_3V_4 \rightarrow V_4C_3V_2C_1-C_1V_2C_3V_4$). Although it remains possible that a linguistic theory could require both pieces and processes (such as a re-write rule being triggered by a past tense morpheme in certain English words—e.g. $\text{sing} \rightarrow \text{sang}$), I take it to be the case that a piece-based framework is to be preferred over a process-based framework, if it is tenable. This is especially the case since the piece-based view gives us something to compare when we do cross-linguistic investigation into closely-related languages. (This issue will be particularly relevant in Part II on reduplication across Uto-Aztecan).

The second point to be decided upon in theory-selection is what it is that we consider to be manipulated by the syntax. One current approach advocates what has been termed the "Lexicalist Hypothesis" (after Chomsky 1970), which essentially holds that some (or
all) “words” are created “in the Lexicon” for insertion into syntactic terminal nodes. For example, Chomsky’s ungrammatical example of *John’s growth of tomatoes was taken to be proof that a causative verb grow had to be created in syntax, since a generative lexicon should be able to produce such an example (cf. The army’s destruction of the city, where destruction was taken to be derived in the lexicon itself). In Marantz (1997a’s view, which is the point of view adopted by most practitioners of DM, roots are category-neutral, and noun-hood and verb-hood are determined by the syntactic context in which roots are inserted (cf. also Harley and Noyer 1998). The ungrammaticality of a causative nominalization of an inchoative verb like √GROW is explained by the fact that such roots belong to a class of items which denote internally-caused changes of state, and therefore cannot have external causers. The issue therefore is not that some verbs are derived in the “Lexicon” and others in syntax proper, but by positing that all derivation occurs in syntax we merely note that not all roots are semantically compatible with all syntactic positions.

While Marantz (1997a) maintains that lexicalist models are simply wrong because they make false predictions, these models do provide powerful means for deriving syntactic structures from lexical information. DM, on the other hand, is highly constrained in that it posits a fixed range of available structures, and roots may be inserted into various positions based on their semantic and syntactic features. It should be pointed out, however, that this dissertation is not presented as an effort to debunk one

---

1 The Lexicalist approach is characteristic of several current theoretical models, including Head-Driven Phrase Structure Grammar (HPSG) (Sag and Wasow 1999) and Lexical-Functional Grammar (LFG) (Bresnan 2001).
theory or another, since I presume that both Lexicalist and non-Lexicalist theories could account for the phenomena that I will discuss in later chapters. Crucially, however, Lexicalist theories give privileged status to “words” as elements to be inserted into syntax. In DM, on the other hand, the units to be manipulated by syntax are features, and “words” are an artifact of the insertion of morphophonological units into terminal syntactic nodes, and their subsequent combination in the Morphological and Phonological components of the grammar. It is primarily for this reason that I have adopted DM as a fruitful theory from which to explore comparative morphosyntax in Uto-Aztecan.

Now shifting our focus to DM proper, we should consider the characteristic aspect from which this framework derives its name, distributed, which “refers to the separation of properties which in other theories are collected in the Lexicon” (Harley and Noyer 1999: 3). In DM, there is no primitive notion of a “word”; hence, Marantz (1997b) has famously claimed that the English lexical item cat is a “phrasal idiom”. Since there is no notion of “word” per se, we must consider morpho-phonological “pieces” to be Vocabulary Items (VI’s), which include both roots and affixes.

A schematic diagram of the grammatical levels posited by DM are given in Figure 3.1.
In essence, DM is a classic Y-model, well-established in the generative tradition. In what Harley and Noyer call "List A" are the morphosyntactic features that are manipulated by the operations typical of syntax; in this framework, these operations are, by hypothesis, limited to Merge and Move (i.e. Copy). As with the classic GB model, the grammar in DM splits into two separate modules, Phonological Form (PF) and Logical Form (LF). Crucially, however, there is an intermediate level between syntax and PF,
Morphological Structure (MS). It is here that Vocabulary Items are inserted to give morphophonological realization to the morphosyntactic features that have been arranged by syntax. This is the crucial notion of Late Insertion, "the hypothesis that the phonological expression of syntactic terminals is in all cases provided in the mapping to Phonological Form (PF). In other words, syntactic categories are purely abstract, having no phonological content" (Harley and Noyer 1999: 3).

Another feature of DM is underspecification, which holds that "phonological expressions need not be fully specified for the syntactic positions where they can be inserted... Vocabulary Items are in many instances default signals inserted where no more specific form is available" (3). This underspecification is motivated by the Subset (or Paninian or "Elsewhere") Principle, which is stated in (1):

(1) The Subset Principle (Halle 1997, cited by Harley and Noyer 1999: 5)
The phonological exponent of a Vocabulary Item is inserted into a morpheme... if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary Item contains features not present in the morpheme. Where several Vocabulary Items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal must be chosen.

The final distinguishing feature of DM is what Harley and Noyer refer to as "Syntactic Hierarchical Structure All the Way Down". This idea "entails that elements within syntax and within morphology enter into the same types of constituent structures (such as can be diagrammed through binary branching trees). DM is piece-based in the sense that the elements of both syntax and of morphology are understood as discrete constituents instead of as (the results of) morphophonological processes" (1999: 3). I refer to the theoretical stance taken under each of these anti-Lexicalist assumptions as the "Pervasive
Syntax Perspective” (PSP). This notion is intended to evoke the idea expressed in the title of Marantz’s (1997a) paper: there is “no escape from syntax”.

A theory like DM is necessarily Constructionalist. Vocabulary Items are interpreted in the context(s) in which they appear. Under this view, all form-meaning pairings are idioms; as Harley and Noyer (1999) put it, “the term idiom is used to refer to any expression (even a single word or subpart of a word) whose meaning is not wholly predictable from its morphosyntactic structural description. . . [F]-morphemes [i.e. functional items—JH] are typically not idioms, but l-morphemes [i.e. lexical items—JH] are always idioms” (p. 4). In addition, some meaning is structural in nature, being entirely composed from the meanings of F-morphemes. For example, McGinnis (2002) shows that the structural meaning associated with aspect is both systematic and compositional even in conventional idioms, suggesting that aspctual meaning, even in idioms, is composed in syntax rather than being idiosyncratic to lexical items (or “idiomatic” phrases).

Borer (2003) refers to models of this nature as “exo-skeletal”, since they “share to a varying degree a view of the grammar in which at least some argument structure interpretation is divorced from the lexical entry and rather, is determined by the structure” (p. 32). Endo-skeletal models, including Lexicalist theories, focus on the lexical item and how it projects an argument structure, but exo-skeletal models “focus on the way in which the structure, rather than the listed item, determines not only grammatical properties, but also the ultimate fine-grained meaning of lexical items themselves (an effect at times called coercion)” (p. 33, emphasis in original).
Although this aspect of meaning-interpretation is central to DM, most of our focus in this dissertation will be on the comparison of the actual linguistic pieces (morphemes) themselves, as well as on the syntactic processes that derive complex morphological units ("words").

3.4. Comparative syntax and syntactic reconstruction

The cross-linguistic investigation of syntax can be divided into two distinct areas of inquiry. The first is comparative syntax, the study of which seeks to ascertain the range of possible syntactic variation among languages. This line of investigation is typified by work in linguistic typology. The second is diachronic syntax, which seeks to understand how systems of syntax change over time, and which includes the investigation of both long-range syntactic change as well as the acquisition of new grammars by children (as emphasized especially by such work as Lightfoot 1991). In a study such as the present one, the questions raised by these two approaches are intertwined.

In this section we consider three approaches to these issues: work in grammaticalization theory from the functionalist perspective (3.4.1), a recent approach used to adapt ideas from grammaticalization to the Minimalist Program (3.4.2), and the prospects for doing syntactic reconstruction (3.4.3).

3.4.1. Grammaticalization and Grammaticalization Theory

One of the most prominent current trends in the study of comparative and historical syntax is grammaticalization (or grammaticization or grammatization) theory, as
represented by such works as Heine et al. (1991), Traugott and Heine (1991), Hopper and Traugott (1993), Bybee et al. (1994), and Heine and Kuteva (2002).

The central area of inquiry in grammaticalization studies is how functional categories in language are developed diachronically, typically from lexical material but sometimes also from other grammatical (i.e. functional) forms. This process often involves semantic bleaching and phonological erosion of lexical items as they become more functional and more frequent. Heine (2003: 579) gives the following as the “four interrelated mechanisms” involved in the “grammaticalization of linguistic expressions”:

i. desemanticization (or “bleaching,” semantic reduction): loss in meaning content;

ii. extension (or context generalization): use in new contexts;

iii. decategorialization: loss in morphosyntactic properties characteristic of the source forms, including the loss of independent word status (cliticization, affixation);

iv. erosion (or “phonetic reduction”), that is, loss in phonetic substance.

An array of criticisms of grammaticalization theory has recently appeared in Campbell (2001b). The central points at issue are such questions as whether or not grammaticalization is unidirectional, a claim which states that the development of grammatical categories leads “from concrete/lexical to abstract/grammatical meanings” (Heine 2003: 577), but not vice versa (cf. Campbell 2001a, Janda 2001); whether or not there are other processes involved in historical change than grammaticalization, or if grammaticalization itself is an instantiation of independent processes (e.g. reanalysis) and
should therefore be regarded as epiphenomenal (Campbell 2001a, Joseph 2001, Newmeyer 2001); or whether or not grammaticalization theory is even a “theory” to begin with. It is not the purpose of this dissertation to analyze specific claims that have been propounded by particular grammaticalization theorists; see Campbell (2001b) for the major recent critiques of the approach, and Heine (2003: 581-4) for a response to criticisms of foundational issues. Of more interest to us here are even stricter limitations on grammaticalization that have recently been posited, from comparative Uto-Aztecan evidence, by Heath (1998) and J. Hill (2003).

Heath (1998) presents a firm rebuttal to the tenets of grammaticalization theory, using data from Northern Uto-Aztecan (NUA) languages to show that grammaticalization processes are often not “directed” by “natural paths” of grammatical change, but come about via the fortuitous phonological similarity of “host” morphemes which new elements can glom onto. Heath’s metaphor is that old affixal categories are “hermit crabs” that are successively located in new “shells”—morphemes that take different pre-existing phonological forms and add the new, “grammaticalized” meaning. An example of this is NUA *-nac- ‘sit down’ coming to mean ‘future durative’ in Cahuilla. Heath argues that this occurred not because of some semantic affinity between the action of sitting down (the verb of which, *-nač-, does not connote the aspectual property of duration) and future durative, but instead the coincidental phonological similarity between the morpheme *-nač- and the previously existing Uto-Aztecan morpheme for ‘future’: *-ni.
J. Hill (2003) takes the hermit crab notion one step further, claiming that in addition to morphophonological similarity this kind of grammaticalization also requires morphosyntactic amenability, in the form of an available syntactic slot. Following Jelinek (1998)'s work on transitivity in Yaqui, Hill shows that the loci of subject agreement in Cupeño are limited to only a few functional projections (TranP ~ vP, VoiceP ≈ AgrO, AspP, and TP). Hill's formal syntactic account complements and constrains Heath's hermit crab grammaticalization approach by predicting where grammaticalization can occur, in addition to the types of elements that can be grammaticalized.

Work in grammaticalization theory has inspired much recent research into cross-linguistic and diachronic syntax. However, the theoretical aspects of work in this area go beyond the scope of this dissertation. For our purposes, I will simply refer to "grammaticalization" as the historical process of semantic bleaching and/or phonological erosion of morphological elements, with no further theoretical implications other than the observance that these processes are widely attested in the world's languages, and specific examples of them will be discussed below. In most cases, such as the loss of lexical word-hood and subsequent development into clitics and, in some cases, even further development into bound affixes, these processes can be regarded as one of the distinct mechanisms of syntactic change addressed by Harris and Campbell (1995), to be discussed in section 3.4.3.
3.4.2. Grammaticalization and Minimalism—Roberts and Rousseau (2003)

One of the central areas of inquiry in modern generative grammar, including Principles and Parameters theory and, more recently, the Minimalist Program (Chomsky 1995 et seq.), is the search for an explanation of cross-linguistic variation in syntax given the underlying assumption that all syntax is constrained by Universal Grammar. A non-exhaustive list of influential work from this perspective includes Borer (1984), Travis (1984), Baker (1988), Lightfoot (1991), Roberts (1993), Kayne (2000), and the papers collected in such works as van Kemenade and Vincent (1997) and Lightfoot (2002c). Most of these authors have focused on microparametric variation in syntax. Baker (2001) presents the most thorough discussion of macroparameters; see Chapter 9 for a critique of this work. In this section I will review a recent proposal to capture the generalizations pointed out by grammaticalization theorists within the Minimalist Program, developed by Roberts and Rousseau (2003).

Working within the Bare Phrase Structure theory of Chomsky (1995), Roberts and Rousseau present an account of parameter-setting that seeks to explain grammaticalization, construed as the reanalysis of lexical or functional material as functional heads, often resulting from the simplification of more complex underlying representations. The relevant functional categories that they consider are I (Inflection), C (Complementizer), and D (Determiner). In Roberts and Rousseau's theory of variation, these syntactic categories may either contain a PF representation (i.e. an overt morpheme that is inserted into that syntactic slot), or they may be marked by a diacritic that serves as a trigger for another element to move into that slot, as with the EPP feature of T that
triggers movement of the subject to that position in English finite clauses. These features are arbitrary, and parameter-setting occurs when feature values are assigned during language acquisition. According to Roberts and Rousseau, “language change consists of some change in the realization/attraction property of functional heads, that is, a change in the lexicon” (p. 7). This view of the “lexicon” corresponds with List A in Figure 3.1. Roberts and Rousseau instantiate parametric change by assigning (randomly) a diacritic (*) “to features typically associated with functional heads. Where the diacritic is assigned to a feature, that feature, F*, must have a PF realization” (p. 29). This * “is assigned in the lexicon, following Borer’s (1984) idea that parametric variation is a facet of the lexicon” (p. 29).

Functional features may be realized in one of two ways—Merge or Move (i.e. Copy). Under Roberts and Rousseau’s view, “Merge is always preferred over Move” (p. 29). This has to do with the nature of the process of acquiring parameters. Roberts and Rousseau postulate that “the learning device is computationally conservative in that it has a built-in preference for relatively simple representations. In other words, if the trigger is ambiguous, the learner will choose the option that yields the simpler representation” (p. 15). For example, under the assumption that movement creates adjunctions (cf. Kayne 1994), Roberts and Rousseau point out that “movement always creates relatively complex representations, in the obvious sense that [2]b with F adjoined to G is a more complex structure that [2]a, where no movement, and thus no adjunction, has taken place” (p. 16):
A shift from (2)b to (2)a can occur under more than one circumstance. First, movement of F to G will not occur unless it is properly triggered by some feature of G. If G loses this feature then movement will not occur, a change in a movement parameter. Second, F can be reanalyzed as a part of G, as in the case of the “misanalysis” of some lexical item as an inflectional element, thus creating new functional material. Another reanalysis involves a complex XP being reanalyzed as a simple X category, thus losing any internal structure (p. 16). Roberts and Rousseau give many examples of parametric changes of this type for I, T, and D elements, from a wide variety of languages. It is the notion of structural simplification of underlying syntactic representations, along with markedness of certain parametric values, that relates Roberts and Rousseau’s Minimalist theory to the empirical results derived from research within grammaticalization theory.

Roberts and Rousseau’s theory is primarily a theory of syntactic change. Most of the comparative Uto-Aztecan syntax that will be discussed in subsequent chapters is morphological in nature, and in most cases what will be discussed is common to most Uto-Aztecan languages and thus my primary concern will be with (morpho)syntactic reconstruction (see section 3.4.3 for discussion of this notion). However, it is interesting to note that what I reconstruct in Part III is based on the notion of movement (rather than base-generation, or Merge), and if my reconstructions are correct, and if my account of the diachronic development of polysynthesis in Chapter 9 is right, then Movement
operations can be maintained without structural simplification for extremely long periods of time; in the Uto-Aztecan case they have apparently been retained for a few thousand years.

3.4.3. On syntactic reconstruction—Harris and Campbell (1995)

Harris and Campbell (1995) develop a general theory of syntactic change, one of the key foci of which is syntactic reconstruction. For Harris and Campbell, there are only three mechanisms of syntactic change: reanalysis, extension, and borrowing. According to their approach, all syntactic change, including grammaticalization, can be recast in these terms.

Following the definition of the term proposed by Langacker (1977b), Harris and Campbell define reanalysis as "a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any modification of its surface manifestation" (p. 50). This "underlying structure" can include constituency, hierarchical structure, category labels, grammatical relations, and "cohesion". This restructuring "depends upon a pattern characterized by surface ambiguity or the possibility of more than one analysis" (p. 51). Reanalysis is the most important type of syntactic change.

An example of a reanalysis of case-marking in the history of Uto-Aztecan was discussed in 2.2.2, wherein the PUA combination of two nominal suffixes, *-ti ‘non-possessed noun’ + *-a ‘ACC’, was reanalyzed as a single morpheme: the accusative case-marker –ta, in some Uto-Aztecan languages (e.g. Hopi and Yaqui). Another example, involving change in cohesion ("the status of a linguistic sequence as a fully independent
word, a clitic, or an affix, or an analyzable part of a larger unit”) (Harris and Campbell 1995: 63), is found in Nahuatl. Here, an older independent word *nemi* ‘to live, to walk’ has lost its word-hood, and has been reanalyzed in different ways in different dialects. In the Nahuatl of Tetelcingo, Michoacan, and North Puebla, it has become a verbal clitic with the meaning of ‘to go around doing’. In Huasteca Nahuatl, this clitic has been further reanalyzed as a habitual marker, and has shifted to the “pre-verb root morphological position occupied by directional morphemes” (p. 64). In Pipil, on the other hand, “*nemi* is still an independent word, but has been grammaticalized as a copula (especially meaning ‘to be located somewhere’) and as the auxiliary in progressives” (p. 392).

Harris and Campbell’s term *extension* is similar to analogy. It denotes “a mechanism which results in changes in the surface manifestation of a pattern and which does not involve immediate or intrinsic modification of underlying structure” (p. 51), thus it “operates to change the syntax of a language by generalizing a rule” (p. 97). Examples of extension include the regularization of case-marking patterns in certain languages, e.g., in the regularization of a single rule to both classes of verbs in Laz, whereas other languages (e.g. Georgian and Svan) retain the complex case-marking rules for the two series of verbs of Common-Georgian-Zan (pp. 100-101). This type of change will play little role in the ensuing discussion of comparative Uto-Aztecan morphosyntax.

Harris and Campbell “use the term *borrowing* to mean a mechanism of change in which a replication of the syntactic pattern is incorporated into the borrowing language through the influence of a host pattern found in a contact language” (p.51). Because of
the centuries of cultural and linguistic domination of Uto-Aztecan speaking peoples by Spanish and English in the post-colonial context of the contact between these languages, borrowing is particularly germane to the Uto-Aztecan languages. Recent studies of the impact of borrowing from Spanish into Nahuatl include Hill and Hill (2004)'s discussion of the introduction of Spanish prepositions into Nahuatl, and Flores Farfán (2004)'s discussion of the shift from a polysynthetic to analytic type in the more Hispanized varieties of Nahuatl. Classic studies of the structural influence of Spanish borrowings in Uto-Aztecan include Hill and Hill (1986)'s discussion of Spanish influence on the Nahuatl spoken in the Malinche Volcano region, and Campbell (1987)'s discussion of Pipil's borrowing of Spanish coordinate conjunctions.

As alluded to above, one of the purposes of Harris and Campbell's approach to historical syntax is the reconstruction of syntactic patterns. The procedure involved in this reconstruction is the comparative method, and what is required to apply this procedure is the identification of syntactic correspondences.

Lightfoot (2002a, 2000b) gives a negative view on the idea of syntactic reconstruction, taking the strong position that there cannot be a sensible notion of correspondence at the sentence level and therefore nothing for the comparative method to compare in syntax. Lightfoot does maintain that "one can reconstruct syntactic patterns where the daughter languages show identity; and sometimes one can reconstruct identity", thus enabling us to "make a guess about the prehistory of an individual language: if a language Anglish, lacking property $p$, is related to a set of languages all manifesting property $p$, we may guess that prehistoric Anglish had that property and lost
it” (p. 120). However, Lightfoot goes on to suggest that “there is no secure basis for that reasoning, but the guess might turn out to be productive and might help us to understand other properties of Anglish, which we might come to see as fossils, remnants of a prehistoric system” (p.120).

In his critique of such work as that of Harris and Campbell (1995), Lightfoot focuses almost exclusively on issues of word order change, and the problems encountered in the use of directionality of change often attributed to the work of linguistic typologists in the 1970s, including such Uto-Aztecanist representatives as Langacker (1977a) and Steele (1975, 1976, 1977, 1979) (see Chapter 2 for discussion of these works).

In their defense, Campbell and Harris (2002) point out that their (1995) reconstructions focus not on corresponding sentences, but on corresponding syntactic patterns. They state that,

Cognates of any kind are related by virtue of descent from a common ancestor; cognate words, for example, are descended from the same word in the protolanguage. Cognate sentences cannot, of course, be descended from the a shared sentence (except in formulaic language, as in legal codes, proverbs, etc.); they are examples of shared patterns descended from a pattern in the proto-language. This is the reason why we emphasized in our [1995] work that our goal is to reconstruct patterns, not sentences. (pp. 606-7, emphasis added)

Second, Campbell and Harris deny that identity is a necessary criterion for syntactic reconstruction, noting that phonological correspondences are typically not ones of identity among the
daughter languages, where the choice of what to reconstruct in the parent language depends on how many languages exhibit the same reflexes, on our knowledge of phonological change, and in particular, on which sounds commonly change into others and which do not. This is as true of morphosyntactic comparisons as it is of phonological or lexical ones. This is standard historical linguistic procedure, where we try to get as full a story as possible and we employ the principles of historical linguistics to arrive at a reconstruction and to explain changes any of the daughter languages have undergone (p. 609).

Moreover, they invoke the principle of economy in explaining the word order changes in Indo-European addressed by Lightfoot:

... it is highly unlikely, if SOV were not the word order [of Proto-Indo-European], that so many of the daughters should have changed their different word orders to SOV independently (Hittite, the first to branch off the family tree, is SOV, as are a number of other early branches—Indic, Hellenic, Italic, Germanic). All this suggests a reconstruction of SOV for the proto-language. The point is not whether the SOV reconstruction for Indo-European is correct, but rather that there are additional considerations beyond the facts of non-identity cited by Lightfoot which may aid in reconstruction within the comparative method. (p. 611)

Campbell and Harris conclude by pointing out that “if languages are related, they are related as wholes—phonology, morphology, syntax, etc.” (p. 616), a notion which echoes the claims of Thomason and Kaufman (1988).
If syntax were no more than word order then Lightfoot’s skepticism regarding successful reconstruction might indeed be well-founded. However, if one accepts that morphology can be reconstructed, as Lightfoot apparently does, then this position begs the question of what the relationship of morphology is to other components of grammar, including syntax. The central areas of inquiry in this dissertation on comparative Uto-Aztecan involve morphosyntactic patterns, which typically instantiate some aspect of morphology as well as syntax. As Harris and Campbell (1995) put it,

Morphology is concrete and phonologically endowed and for this reason is widely regarded as lending itself to reconstruction. In those languages with rich morphology, there is typically a close relationship between the morphology and the syntax. To the extent that the morphology can be reconstructed by the comparative method, many aspects of proto-syntax will become clear. That is, normal techniques of lexical reconstruction, based on the sequence of recurring sound correspondences in cognate words, can be used to reconstruct polymorphemic words. Morphological analysis of these reconstructed proto-words provides the morphology free, as it were, so long as cognate morphemes have not undergone substantial functional or positional shifts. (pp. 358-9).

In some cases, e.g. with the various denominal verb morphemes to be discussed in Chapter 8, the comparative method allows us to reconstruct a proto form of the various morphemes for PUA. In examining the languages and noting the prevalence of modifier “stranding” and hyponomous arguments (Chapters 6-8), we can invoke Hoenigswald
(1991)'s application of the principle of uniformitarianism in historical linguistics to reconstruct these patterns for PUA. As Hoenigswald puts it,

Uniformitarianism is the refusal to posit for the reconstructed past properties and processes different from those which can be observed at present, or, to put it in another way, the principle that proto-languages should not be expected to differ any more from their descendents than the descendents differ from one another. (p. 25)

With a focus on morphology and with this principle in mind, I follow Harris and Campbell in asserting that certain aspects of syntax are amenable to comparison and reconstruction. Even further, however, I hold that one consequence of a DM-style approach is that the syntactic view of morphology entails that the reconstruction of morphology is the reconstruction of syntax. This is particularly true for languages with rich morphology, as the languages in Uto-Aztecan typically are.

3.5. Summary

In this chapter I have presented background on various approaches to syntactic variation, including the theoretical approaches of grammaticalization theory and Principles and Parameters. In addition, I have suggested that the Pervasive Syntax Perspective adopted by practitioners of Distributed Morphology lends itself to morphosyntactic reconstruction. This concludes Part I of this dissertation. We now turn to our consideration of issues in comparative Uto-Aztecan morphosyntax.
4.1. Introduction: Raimy's Modular-Derivational approach to reduplication

Researchers who have examined the morphology of reduplication within the generative tradition can essentially be divided into two camps: those who view the grammatical process of reduplication as a fulfillment of a morphemic piece (i.e. an affix), and those who view reduplication as a morphological process (i.e. something that occurs within a stem).

Among those who hold the former analysis are Marantz (1982), McCarthy and Prince (1993, 1995) and most other work in Optimality Theory, and Travis (1999, 2001). Within this group we can also include proponents of grammaticalization theory, such as Bybee, Perkins and Pagliuca (1994), as well as such traditional work in comparative Uto-Aztecan as Langacker (1977a). Work on reduplication done within process-based theories include Anderson (1992), and Aronoff (1976) gives an account whereby reduplication occurs as a phonological rule.

Stonham (1994) gives an analysis in which reduplication is treated as a process triggered by certain suffixes that impose an output template on the word onto which they are affixed. He gives data from Nitinaht (Wakashan) that shows that suffixes fall into

---

* An earlier version of this chapter was presented at the Eighth Southwest Workshop on Optimality Theory (SWOT 8) at the University of Arizona on April 26, 2003. I thank the audience at that conference for helpful questions and comments.
classes based on the reduplication patterns and effects on the root that they trigger when they are affixed. As Stonham shows, the effects that are triggered on the root by these suffixes can occur even when additional suffixal material intervenes, thus violating adjacency between the reduplicant and its base.

Raimy (2000) is interesting because he gives a process-based analysis of reduplication within the piece-based theory of Distributed Morphology. Raimy's analysis of reduplication involves a re-adjustment rule triggered by a null morpheme. While such an approach is completely consistent with the tenets of DM, which uses such re-adjustment rules in other cases of non-agglutinative affixation (e.g. English run → ran in the context of the feature [+past]), it would be interesting to at least attempt to regard reduplicative morphemes as pieces in their own right in order to determine the limitations of such an analysis.

Raimy's theory presents a novel approach to reduplication wherein the linearization of a phonological string is regarded as a phonological rule that can be ordered with respect to other phonological rules. Reduplication is brought about by a looping mechanism inserted into the linear structure. For example, in (1), the reduplicated form *bukubuku* is derived via a rule that places a loop from the final vowel of the string back to the first consonant (in Raimy's notation, # is the beginning of a string, and % is the end of the string):

---

1 Raimy's analysis is similar to Stonham (1994)'s in this respect, except that in Wakashan the triggering morphemes are actual suffixes with phonological content.
2 The reduplicant appears in bold throughout this dissertation.
From the underlying form \( # \rightarrow b \rightarrow u \rightarrow k \rightarrow u \rightarrow \% \), modified with this looping mechanism, the rule of Linearization creates \( # \rightarrow b \rightarrow u \rightarrow k \rightarrow u \rightarrow b \rightarrow u \rightarrow k \rightarrow u \rightarrow \% \). (Infinite loops are ruled out by economy—once a loop is followed it is “satisfied”).

There are three crucial aspects of Raimy’s theory that I want to note here. The first is that Raimy’s theory abandons all reference to prosody. Templatic effects are brought about by placing the looping mechanism at specific increments within the underlying string. In simple cases, such as core syllable (i.e. CV) reduplication, this could be a rule indicating a loop at the first vowel; in more complicated cases, the location of the link is entirely stipulative. I will delay an extended discussion of Raimy’s approach until section 5, before which I will present the comparative Uto-Aztecan data and my own prosody-based analysis of it.

The second important aspect of Raimy’s theory is that it is derivational, and Raimy specifically argues against parallel processing models such as Optimality Theory (OT) (Prince and Smolensky 1993, McCarthy and Prince 1993, 1995, and much other work). Finally, and most crucially for the DM model, Raimy proposes that reduplication itself is best understood as a re-write rule triggered by a zero morpheme. Thus, reduplicants are not morphemes per se, but are only epiphenomena of the readjustment loop triggered by a null morpheme.

In this chapter I argue that each of these major points are unnecessary in a piece-based theory of morphology. First, prosodic organization is crucial to a full understanding of
the cognate reduplication patterns in Uto-Aztecan. Second, a constraint-ranking framework (such as OT) best captures the cross-linguistic generalizations seen in Uto-Aztecan, since independently attested constraints account for the reduplication patterns that we find, in addition to accounting for other aspects of the comparative phonology of these languages. Third, the historical stability of reduplicative morphemes in Uto-Aztecan indicates that reduplicants themselves are morphemes, as is assumed by work in the OT tradition, and as is suggested by proponents of grammaticalization theory as well.

In sum, I argue that the Uto-Aztecan languages share recurring cognate patterns of reduplication which vary in ways best explained by appealing to both prosodic structure and the to relative ranking of markedness and faithfulness constraints, and that constraint-ranking captures generalizations that would be less gracefully expressed by the traditional rules of historical linguistics.

4.2. Reduplication in Uto-Aztecan

The Uto-Aztecan languages contain a variety of reduplication patterns, and multiple reduplicative morphemes can be reconstructed for Proto-Uto-Aztecan (PUA). Traditional work in comparative Uto-Aztecan linguistics (e.g. Langacker 1977a and Heath 1978) has discussed these morphemes in terms of rules and in terms of segmental attachment to C- and V- slots. Langacker (1977a) notes that “virtually every Uto-Aztecan language displays verbal reduplication of some kind, and in some cases a variety of patterns” (128), but he did not suggest a definitive reconstruction for the Proto-Uto-Aztecan (PUA) reduplication pattern(s). However, Langacker did point to a variety of reduplication
patterns which surface across the Uto-Aztecan family, without making reference to their prosody:\(^3\)

(2) **Uto-Aztecan Reduplication Patterns**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVCV-</td>
<td>‘iterative’</td>
</tr>
<tr>
<td>CV-</td>
<td>‘distributive’</td>
</tr>
<tr>
<td>CV?</td>
<td>‘punctual’</td>
</tr>
<tr>
<td>CV:</td>
<td></td>
</tr>
<tr>
<td>CV</td>
<td></td>
</tr>
</tbody>
</table>

Langacker suggested that the smaller reduplicants may be reduced from the larger ones, and he also documented the equivalent of what I will refer to as *mora affixation*. That is, Langacker identified consonant gemination (or non-lenition, as in Northern Paiute), vowel-lengthening, and glottalization as processes related to reduplication.

Heath (1978) made specific proposals for three reconstructions of Proto-Northern-Uto-Aztecan (PNUA) reduplication patterns, as given in (3):

(3) **Heath’s reconstructions for PNUA**

i. CVCV:- ‘iterative’
ii. CV:- ‘distributive’
iii. CV?: ‘punctual’

The first of these is what I will call *disyllabic, foot, or full root* reduplication; the second seems to be light syllable reduplication. The third morpheme, a glottal-stop infix, is the NUA reflex of what I reconstruct as a bare mora affix, which usually appears (in SUA and in Numic) as morphological gemination.

\(^3\) Langacker’s list is not exhaustive, since other reduplication patterns are also attested in Uto-Aztecan. For example, Cupéno has suffixal -VC reduplication in such forms as *ku.pup* ‘to be sleepy’ from *kup* ‘sleep’, whereas prefixal CV- reduplication with this root, *ku-kup* means ‘to lie in bed habitually’ (J. Hill in press).
Because of the distribution of reduplicative morphemes across the Uto-Aztecan family, in both NUA and SUA, Haugen (in press a) reconstructs four reduplicative morphemes for PUA:

(4) Proto-Uto-Aztecan reduplication patterns

i. light syllable reduplication:  \( \text{RED} = CV \)

ii. “marked heavy syllable” reduplication:  \( \text{RED} = CV_1X \)  \( (X = \text{Gem, V, L}) \)

iii. full root or disyllabic reduplication:  \( \text{RED} = CVCV \) or \( \sqrt{} \)

iv. mora affixation:
   (\( \text{Gem} = \text{gemination}; \text{L} = \text{laryngeal} \))  \( \text{“RED”}^\text{m} = \mu \)

Because each allomorph of reduplication is independently attested in multiple languages in each major sub-branch of the family, we can suppose that they must have also been present in PUA (Haugen in press a). (The semantics of these reconstructed reduplication patterns will have to be determined by further comparative work, although the distribution of a marked heavy syllable pattern for ‘distributive’ in NUA and SUA makes at least that reconstruction for PUA likely—see discussion below and in Haugen in press a).

For our purposes in this chapter I will focus only on the two reduplicative allomorphs that evince cross-linguistically variable constraint-ranking among these related languages. Both involve “mora augmentation” (Davis 2001): the enhancement of a phonological string with an additional mora for some morphological purpose.\(^5\) These are “marked heavy syllable reduplication” and mora affixation.

\(^4\) The RED morpheme here appears in scare quotes because it is not actually analyzed as a reduplicative morpheme—i.e. a morpheme showing correspondence between the reduplicant and its base. It is merely a prosodic unit, a mora, inserted into the underlying structure. See Samek-Lodovici (1992) and Haugen (2003) for formal treatments of this bare mora morpheme in terms of Optimality Theory.

\(^5\) Davis (2001) makes a distinction between prosodic and morphological mora augmentation. The examples discussed here are of the morphological type.
From a morphophonological perspective that allows for multiple ways of realizing a modification of underlying prosody, i.e. by mora augmentation and the "Emergence of the Unmarked" (TETU) (McCarthy and Prince 1994), there are (at least) five possible ways to realize an extra mora in an output string of segments. These are shown in (5):

(5) Ways to realize the extra mora in marked heavy syllable reduplication:

| (a) continue copying: | noka → nok.noka |
| (b) gemination:       | noka → noo.noka |
| (c) vowel lengthening:| noka → no?.noka |
| (d) epenthesis of an unmarked consonant: | noka → no?.noka |
| (e) "nasal substitution": | papa → pam.papa |

The pattern in (5a), typical of languages like Ilokano (McCarthy and Prince 1986), is the expected pattern, based on the idea that full copy of a base is a primitive requirement of reduplication (e.g. as encapsulated in the OT constraint "MAX-BR", McCarthy and Prince 1995). However, this pattern is unexpectedly atypical in Uto-Aztecan. The usual pattern is that of (5b), where gemination is triggered between the base and reduplicant.

The second most common pattern in Uto-Aztecan is (5d), where an unmarked consonant is epenthesized to express the extra mora. Vowel lengthening (5c) is also attested, but perhaps only in languages that also have gemination or epenthesis. The final pattern, "nasal substitution", is typical of Pohnpeian (Austronesian) (Kennedy 2003), but is not attested in Uto-Aztecan.\(^6\)

\(^6\) This appears to be true at least synchronically. However, there may be some evidence of nasal substitution historically in Tübatalabal—see Crowhurst (1991) for a synchronic description and pre-OT analysis of the interesting interactions of nasal codas in Tübatalabal reduplication.
From the point of view of Prosodic Morphology, where we consider the similar, indeed, the nearly identical patterns of heavy syllable reduplication to be related, we must make reference to the fact that these patterns actually involve a heavy syllable, and further, that these heavy syllable reduplication patterns contrast with light syllable reduplication patterns in the same languages (thus my use of the term marked heavy syllable reduplication), in order to consider them to be cognate patterns. If we follow Raimy and do not consider the prosodic nature of these reduplication patterns, it is unclear in what sense we could consider these patterns to be cognate. Similarly, under the traditional comparative method we would compare the segmental content of these morphemes, rather than refer to the common prosody that they share.

We now turn to the cross-linguistic data that illustrate the cognate status of these morphemes in Uto-Aztecan. For both of the patterns under discussion I will illustrate cognate reduplicants in both NUA and SUA in order to illustrate the wide-spread geographical distribution of these patterns within the family. See Haugen (in press a) for further discussion of the full range of reduplicative allomorphy across Uto-Aztecan, evidence for the reconstruction of each of the four distinct patterns given in (4), as well as for arguments against an areal diffusion account of the distribution of these morphemes.

4.2.1. “Marked heavy syllable reduplication”

The first reduplication pattern that we will discuss is “marked heavy syllable reduplication” (Haugen 2003, in press a). Before discussing this pattern, though, we must first discuss the issue of allomorphy in reduplication more generally. Specifically,
Spaelti (1999) distinguishes between two types of multi-pattern reduplication. The first type, the *dupleme*, is the pattern of variation that we see when two differently shaped reduplicants have different meanings, and the meaning is predictable based on the shape. The second type is the *alloduple*, where differently shaped patterns of reduplication occur based on predictable phonological properties of the base. Marked heavy syllable reduplication is generally a pattern which marks a different morpheme (i.e. encodes a different semantic function) from a light syllable reduplicant in the same language, and is thus an instance of a second *dupleme*. This heavy syllable reduplicant is “marked” in the sense that it is a morpheme that “marks” a light syllable reduplication pattern with an extra mora to create a morpheme that has a meaning distinct from that which is expressed by light syllable reduplication.\(^7\) It is also crosslinguistically “marked”, in the sense that reduplication is generally expected to maximally copy the “base”, but in most UA languages one of the other mechanisms of lengthening is usually preferred over further copying (see 5). That is, the base for reduplication in Uto-Aztecan is usually limited to the first syllable of the stem.

4.2.1.1. Yaqui (SUA)

Yaqui has both kinds of multi-pattern reduplication—duplemes and alloduples, and it therefore has a mixed system with respect to Spaelti’s types of multi-pattern reduplication. The default pattern of reduplication in Yaqui is so-called “syllable copy”

---

\(^7\) In this sense, marked heavy syllable reduplication can be seen as an instance of “enhancing contrast” between reduplicative morphemes, in the sense of Urbanczyk (2002).
reduplication,\(^8\) where the entire first syllable is copied in order to indicate habitual action. If the first syllable of the base is of the shape CV, then the reduplicant is a CV (6); if the first syllable of the base is a CVC syllable, then the reduplicant is (usually) a CVC syllable (7):

(6) Yaqui monosyllabic habitual reduplication: CV base (Molina et al. 1999)

<table>
<thead>
<tr>
<th>a.</th>
<th>vu.sa</th>
<th>‘awaken’</th>
<th>vu.vu.sa</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td>chi.ke</td>
<td>‘comb one’s hair’</td>
<td>chi.chi.ke</td>
</tr>
<tr>
<td>c.</td>
<td>he.wi.te</td>
<td>‘agree’</td>
<td>he.he.wi.te</td>
</tr>
<tr>
<td>d.</td>
<td>ko.’a.rek</td>
<td>‘wear a skirt’</td>
<td>ko.ko.’a.rek</td>
</tr>
<tr>
<td>e.</td>
<td>cho.’i.la</td>
<td>‘lasso’</td>
<td>cho.cho’il.a</td>
</tr>
</tbody>
</table>

(7) Yaqui monosyllabic habitual reduplication: CVC base (Molina et al. 1999)

<table>
<thead>
<tr>
<th>a.</th>
<th>vam.se</th>
<th>‘hurry’</th>
<th>vam.vam.se</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td>chep.ta</td>
<td>‘jump over’</td>
<td>chep.chep.ta</td>
</tr>
<tr>
<td>c.</td>
<td>chuk.ta</td>
<td>‘cut with a knife’</td>
<td>chuk.chuk.ta</td>
</tr>
<tr>
<td>d.</td>
<td>hit.ta</td>
<td>‘make a fire’</td>
<td>hit.hit.ta</td>
</tr>
<tr>
<td>e.</td>
<td>bwal.ko.te</td>
<td>‘soften, smooth’</td>
<td>bwal.bwal.ko.te</td>
</tr>
</tbody>
</table>

There are a few cases where a coda consonant does not copy (e.g. \textit{bwakta} \rightarrow \textit{bwa.bwa.kta} ‘take out of a container’), but no cases where we see copy into a second syllable: i.e. there are no forms like *\textit{vus.vusa}. There are also some cases, within a certain class of words, that maintain a long vowel in the base, and because the reduplicant is a light syllable with a short vowel the full first syllable of these words is also not copied (e.g. \textit{ka.a.te} \rightarrow \textit{ka.kaa.te}, *\textit{kaa.kaa.te} ‘they are walking’)—see further discussion of this class in Demers, Escalante and Jelinek (1999).

\(^8\) “Syllable-copy” reduplication has often been claimed not to exist (cf. Moravcsik 1978, Marantz 1982, McCarthy and Prince 1986, etc.). The Yaqui facts contradict this claim; see Haugen (2003) for discussion of this issue with respect to Yaqui. Ballantyne (1999) also gives evidence for syllable-copy reduplication in Yapese (Micronesian).
Because the reduplicant always surfaces with a short vowel, Demers et al. (1999) and Haugen (2003) claim that the reduplicant is a light syllable, and that coda consonants are not moraic, except when they are geminate consonants serving some morphological function (i.e. mora affixation or marked heavy syllable reduplication). Also, Haugen (2003) argues that the base for reduplication in these cases is only the first syllable of the word.

Harley and Amarillas (2003) show that there are multiple phonological patterns of reduplication in Yaqui, and that each can serve a variety of semantic functions. Haugen (2003) analyzes the phonological patterns specific to habitual reduplication: “syllable copy” (i.e. light syllable reduplication); CVCV- reduplication, and mora affixation, and argues that the root-specific allomorph is not predictable from the underlying phonological form. That is, there must be some lexical specification as to which root takes which alloduple.® Thus, there are some near minimal pairs for the Yaqui habitual (data from Molina et al. 1999):

(8)a. ívakta 'embrace' → i.'i.vak.ta
b. kinakta 'squint, grimace' → ki.na.ki.nak.ta
c. máveta 'receive' → mov.ve.ta

This lack of predictability between reduplicant form and meaning is not limited to Yaqui within Uto-Aztecan. Tuggy (2003) points out that reduplicant shape and meaning

---

® This claim is consistent with Kager (1999)’s discussion of reduplication in Southeastern Tepehuan (Tepiman), where he comments on the variation in light syllable and heavy syllable reduplication by stating that “the prosodic size of the reduplicative prefix (a heavy syllable or light syllable) is based on a stem-specific choice... ‘reduplicative allomorphy’ is apparently not governed by phonological or morphological factors” (182). See section 4.4 for discussion of similar long and short vowel reduplication in another Tepiman language, Tohono O’odham.
correspondences are also not at all transparent in Nahuatl, either synchronically or
diachronically. Tuggy points to Carochi’s oft-cited quotation from 1645:

To know upon what occasions this first syllable is to be doubled, and how it is to
be pronounced, whether with a saltillo [i.e. glottal stop or [h]—JH], or
without it, and to know, precisely what the verb means, when the doubled first
syllable has a saltillo, and when it has the long accent, is the most difficult thing
that there is in this language, and I doubt that those who do not know it naturally,
can possibly conquer this difficulty. (Carochi 1645; cited and translated by
Tuggy 2003: 127)

Returning to the issue of the non-habitual dupleme in Yaqui, the marked heavy
syllable reduplicant surfaces as a heavy syllable that triggers gemination from the onset
of the base into the coda position of the reduplicant.\(^\text{10}\) The semantics of such
reduplication is usually a kind of iterative indicating ‘from time to time’ (Demers et al.
1999), although there is variation between speakers and there are many roots which have
an idiosyncratic meaning with such reduplication. Some examples of this pattern are
given in (9):

(9) **Yaqui marked heavy syllable reduplication** (Molina et al. 1999)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>bwii.ka</td>
<td>‘sing’</td>
</tr>
<tr>
<td>b</td>
<td>tee.ka</td>
<td>‘lay it across’</td>
</tr>
<tr>
<td>c</td>
<td>va.hu.me</td>
<td>‘swim’</td>
</tr>
<tr>
<td>d</td>
<td>ye.na</td>
<td>‘smoke (tobacco)’</td>
</tr>
<tr>
<td>e</td>
<td>‘om.te</td>
<td>‘get angry’</td>
</tr>
</tbody>
</table>

\(^{10}\) Previous literature has referred to this pattern of reduplication in Yaqui as “secondary reduplication” (e.g.
Escalante 1985, Demers et al. 1999, Molina et al. 1999, etc.). This term has also been applied to the
cognate pattern of reduplication in related Uto-Aztecan languages of Miller’s proposed “Sonoran” branch
of SUA, in Barragan and Haugen (2002), the first effort at connecting the Yaqui pattern with other
“secondarily heavy” syllable reduplication patterns in Uto-Aztecan.
We now turn to the cognate heavy syllable reduplication patterns in other Uto-Aztecan languages.

4.2.1.2. Mayo (SUA)

Yaqui’s closest linguistic relative is Mayo, the other language in the Cahitan sub-group of Taracahitic. Mayo has cognates for both the light syllable reduplicant and the marked heavy syllable reduplicant of Yaqui. However, Hagberg (1993) reports that the semantic distinction between the two has been lost in all but the eldest generation of speakers. Thus, what remain duplemes in the elders’ speech have merged into alloduples of a single dupleme for younger speakers, in that there is free variation between light and heavy syllable reduplication for a single semantic function: the expression of habitual action.

Hagberg (1993) divides the Mayo lexicon into two classes. The first is the *accented* class, with accent on the first vowel (10), and the *unaccented* class, with accent on the second vowel (11). These two classes differ with respect to what is copied in heavy syllable reduplication (“RED2” in 10 and 11):

(10) **Mayo accented words** (Hagberg 1993)

<table>
<thead>
<tr>
<th>Stem</th>
<th>RED1 = σ₁</th>
<th>RED2 = σ₁₀₀</th>
<th>Unattested</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. yú.ke</td>
<td>yú.yu.ke</td>
<td>yúy.yu.ke *yúk.yu.ke</td>
<td>'rain'</td>
<td></td>
</tr>
<tr>
<td>b. wóm.te</td>
<td>wó.wom.te</td>
<td>wów.wom.te *wów.wom.te</td>
<td>'be frightened'</td>
<td></td>
</tr>
<tr>
<td>c. nók.wa</td>
<td>nó.nok.wa</td>
<td>nók.nok.wa *nón.nok.wa</td>
<td>'known language'</td>
<td></td>
</tr>
<tr>
<td>d. nó.ká</td>
<td>nó.no.ká</td>
<td>nó.n.o.ká *nók.n.o.ká</td>
<td>'know a language'</td>
<td></td>
</tr>
</tbody>
</table>
(11) **Mayo unaccented words** (Hagberg 1993)

<table>
<thead>
<tr>
<th>Stem</th>
<th>RED1=σ₁</th>
<th>RED2=σ₂</th>
<th>Unattested</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bwa.ná</td>
<td>bwa.bwa.ná</td>
<td>bwan.bwa.ná</td>
<td>*bwa.bwa.ná</td>
<td>'cry'</td>
</tr>
<tr>
<td>b. bwi.ká</td>
<td>bwi.bwi.ká</td>
<td>bwik.bwi.ká</td>
<td>*bwib.bwi.ká</td>
<td>'sing'</td>
</tr>
<tr>
<td>c. om.té</td>
<td>om.om.te</td>
<td>om.om.te</td>
<td>*o'.om.te</td>
<td>'hate'</td>
</tr>
<tr>
<td>d. no.ká</td>
<td>no.nó.ká</td>
<td>nok.nó.ká</td>
<td>*non.nó.ká</td>
<td>'speak'</td>
</tr>
</tbody>
</table>

The contrast between the heavy syllable reduplicant (RED2) in (10d), *nón.no.ka*, and (11d), *nok.nó.ka*, shows that the base for reduplication must be different in the two classes of words (Hagberg 1993, Haugen 2004). The copying of the onset of the second syllable of the base evident in Mayo unaccented forms like *nok.nó.ka* is one of what appears to be a very few examples of copying out of the first syllable within Uto-Aztecan.

Although the details of Mayo reduplication are somewhat distinct from those Yaqui, it is clear that the patterns here are cognate with those of Yaqui. Both languages have light and heavy syllable reduplicants. Whereas Yaqui heavy syllable reduplicants are only formed with gemination of the onset of the base into coda position of the reduplicant, Mayo allows either this or copy into the second syllable of the stem, depending on underlying stress and therefore depending on word class. This latter pattern is what is expected from cross-linguistic examination of reduplication patterns, but it is rare and not typical of heavy syllable reduplication in Uto-Aztecan.

4.2.1.3. Guarijio (SUA)

Like many Uto-Aztecan languages, Guarijio nouns have special morphology for human
Plurals in Guarijí'o are normally not marked at all on nouns, but human nouns reduplicate, usually at least optionally with a heavy syllable. In the case of Guarijí'o, however, these marked plurals show up not with gemination or a long vowel, but with an epenthesized laryngeal (usually a glottal stop):

(12) Guarijí'o “marked heavy syllable” = “marked plural” (Miller 1996a)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>ku.ci.ta</td>
<td>kuʔ.ku.ci</td>
</tr>
<tr>
<td>b.</td>
<td>ma.la.la</td>
<td>maʔ.ma.la.la</td>
</tr>
<tr>
<td>c.</td>
<td>se.pu.ri</td>
<td>seʔ.se.pu.ri</td>
</tr>
<tr>
<td>d.</td>
<td>pa.mi.la</td>
<td>paʔ.pa.mi.la</td>
</tr>
<tr>
<td>e.</td>
<td>no.la</td>
<td>noʔ.no.la</td>
</tr>
</tbody>
</table>

‘son, daughter’
‘daughter’
‘uncle, aunt’
‘boss, govenor’
‘son’

Haugen (in press a) alludes to the possibility that these marked plurals are the remnants of a more complicated system, wherein non-human nouns were once reduplicated with a light syllable for the plural in contrast to this (synchronically optional) heavy syllable reduplication for human nouns, a system reminiscent of that in Tepiman (see section 4.4). However, the light syllable reduplication may have been lost once verbal reduplication began to function as plural agreement.

Guarijí'o also has a heavy syllable in distributive reduplication (possibly only in a few examples with numerals), also with epenthesis of glottal stop:

(13) Guarijí'o heavy syllable distributive reduplication (Miller 1996a)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>naó</td>
<td>naʔ-naó</td>
</tr>
<tr>
<td>b.</td>
<td>mariki</td>
<td>maʔ-máriki</td>
</tr>
</tbody>
</table>

‘cuatro/four’
‘de cuatro en cuatro’ / ‘four by four’
‘cinco/five’
‘de cinco en cinco’ / ‘five by five’

Miller (1996a) also alludes to a few other possible examples of a contrasting heavy syllable reduplication pattern, but he notes that his information was sketchy and that a

---

more systematic effort would have to be made to collect such forms (see also Barragan and Haugen 2002).

4.2.1.4. Nahuatl (SUA)

Nahuatl also utilizes multiple patterns of reduplication that are instantiations of “marked heavy syllable reduplication”. Canger (1981) assesses data from various dialects of Nahuatl and argues for reconstructing at least three reduplication patterns for Proto-Aztecan. The first is reduplication with a short vowel, the other two are CV?-reduplication for ‘distributive’ and CV:- reduplication for ‘consecutive’. Canger shows that these patterns must have been productive in Proto-Aztecan, and the evidence considered in Haugen (in press a) and in this chapter suggests further that these patterns are the remnants of the marked heavy syllable reduplication pattern that was productive in Proto-Uto-Aztecan itself.

As in Guarijío, in many dialects of Nahuatl laryngeal-insertion is used to indicate distributive numerals (at least in the “low numbers”, Canger 1981: 36). In the Nahuatl spoken in Xalatzala (Central Guerrero), this laryngeal is typically [h], although in other dialects it can also be a glottal stop (data from Canger 1981):

(14)a. oh-o:me  ‘two at a time’
b. yeh-ye:yi  ‘three at a time’
c. nah-na:wi  ‘four at a time’
d. ma:-ma:k^ili  ‘five at a time’

\[12\] Nahuatl dialects vary as to whether the “saltillo” is expressed as a glottal stop or an [h]; what is crucial here is that since both are laryngeals with unspecified place features, both are amenable to epenthesis. I assume that they are related and which actually surfaces is a dialect-specific choice.
Nahuatl also uses CVL (where L = 'laryngeal') reduplication for expressing distributive action in verbs:

Nahuatl reduplication for distributive action (Canger 1981: 37)

(15)a. ni-k-te:-teki \( \rightarrow \) in tlaškalli  
    I-it-RED-cut \quad \text{the bread}  
    'I cut/slice the bread'  

b. ni-k-\( \text{t} \)-teki \( \rightarrow \) in tlaškalli  
    I-it-RED-cut \quad \text{the bread}  
    'I cut the bread to pieces'  

(16)a. ni-k-\( \text{š} \)-te:-loa \( \rightarrow \) in tlaolli  
    I-it-RED-spread \quad \text{the corn}  
    'I spread the corn out'  
    (over a continuous area)  

b. ni-k-\( \text{š} \)-te:-loa \( \rightarrow \) in tlaolli  
    I-it-RED-spread \quad \text{the corn}  
    'I spread the corn out in different places'  
    (from different piles)  

To conclude this section on marked heavy syllable reduplication in SUA, we have seen heavy syllable reduplication that triggers gemination (Yaqui, the Mayo accented class), epenthesis of a laryngeal (Guarijío, Nahuatl), and copy into the second syllable (the Mayo unaccented class). We will now consider similar patterns in NUA.

4.2.1.5. Numic (NUA)

Heavy syllable reduplicants also appear in Northern Uto-Aztecan (NUA), in the context of plural and/or distributive nouns, in at least Comanche and Tūmpisa Shoshone, both of which are within the Numic branch of NUA. However, in these cases the process does not seem to be productive, and the heavy syllable reduplicants may be fossilized (examples cited in Hill and Hill 2000):

(17) **Comanche number marking** (Charney 1993)

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nam.mi</td>
<td>na.na.na.mi ( \rightarrow )</td>
<td>'sister'</td>
</tr>
<tr>
<td>b. pi.a</td>
<td>pi.pi.a [pi.via] ( \rightarrow )</td>
<td>'big group'</td>
</tr>
<tr>
<td>c. ta.ka</td>
<td>tah.ta.ka.nii ( \rightarrow )</td>
<td>'relatives'</td>
</tr>
<tr>
<td>d. ten.( \text{s} )éé</td>
<td>tēh.ten.( \text{s} )éé ( \rightarrow )</td>
<td>'ten cents apiece'</td>
</tr>
<tr>
<td>e. pi.e.ti</td>
<td>pih.pie.ti.nii ( \rightarrow )</td>
<td>'group of elders'</td>
</tr>
</tbody>
</table>
From (17)d and e it looks like the heavy syllable reduplicant in Comanche indicates distributivity. Human plurals in Tümpisa Shoshone also yield heavy syllable reduplicants:

(18) Tümpisa Shoshone number marking (Dayley 1989)

<table>
<thead>
<tr>
<th></th>
<th>Tümpisa Shoshone number marking (Dayley 1989)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>pa.pi ‘older brother’ pa.pi.am.mü ‘older bro. (pl.)’</td>
</tr>
<tr>
<td>b.</td>
<td>pat.si ‘older sister’ pap.pat.si.am.mü ‘older sister (pl)’</td>
</tr>
<tr>
<td>c.</td>
<td>pe.tü ‘daughter’ pep.pe.tü.mü ‘daughter (pl)’</td>
</tr>
<tr>
<td>d.</td>
<td>tangum.mü ‘man’ tat.tangum.ku ‘man (dl)’</td>
</tr>
<tr>
<td>e.</td>
<td>tokkwapii ‘aunt’ tot.tok.kwapamlü ‘aunt (pl.)’</td>
</tr>
<tr>
<td>f.</td>
<td>tua ‘son’ tut.tu.am.mü ‘son (pl.)’</td>
</tr>
</tbody>
</table>

In Comanche, the heaviness of the syllable is indicated by means of laryngeal- (i.e. [h]-) insertion, rather than fully copying a coda consonant from the base (e.g. ten.see → téh.ten.see and not *tén.ten.see). As with many other UA languages, Tümpisa Shoshone heavy syllable reduplication appears with gemination from the onset of the base to the coda of the reduplicant. While it remains the case that these examples of heavy syllable reduplicants might not be productive in the synchronic state of the language, these forms at least suggest that a marked heavy syllable pattern was once productive in the Numic languages (i.e. marked heavy syllable reduplication is now fossilized).

With a marked heavy syllable reduplicant indicating distributive in at least Guaríjio, Nahuatl, Tohono O’odham, and Comanche, it seems plausible to reconstruct this for PUA. In contrast, Heath (1978) reconstructs CV for ‘distributive’ in NUA. Similarly, Langacker (1977a) states that “distributive value would not be implausible” for CVCV reduplication and its reduced forms: “CVC, CV’-, and possibly even CV:- on stems with an underlying short vowel)” (p. 129). Once we take the prosodic approach we can see
that the cognate reduplicant is similar to what Langacker proposed, in that the same reduplicant, a heavy syllable, surfaces in multiple languages.

However, the cognate status of this distributive morpheme only makes sense when we consider it to be a specific prosodic unit in each of the languages, even though the phonological processes that realize this morpheme may be different in each language. I will argue in section 4.3 that these differences are best accounted for by reference to language-specific ranking of universal constraints (i.e. Optimality Theory).

4.2.2. Mora affixation

We turn now to the second type of morphological mora augmentation found in Uto-Aztecan: mora affixation. Following Samek-Lodovici (1992), Haugen (2003, in press a) posits a bare mora affix as a morpheme in Yaqui and other Uto-Aztecan languages. Depending on the phonology of a given language, this moraic morpheme, which is essentially a bare mora inserted into the phonological string of the stem to which it attaches, can manifest itself by means of morphological gemination, vowel-lengthening, or epenthesis of an unmarked consonant (e.g. a laryngeal). The constraint-based approach predicts that this realization should be the same for mora affixation as it is for marked heavy syllable reduplication.

4.2.2.1. Mora Affixation in Yaqui (SUA)

One of the alloduples of the Yaqui habitual is morphological gemination:
Yaqui habitual mora affixation (Molina et al. 1999)

a. b^a.ta.ni.a ‘burn (food)’ b^at.ta.ni.a
b. e.ta.po ‘open up’ et.ta.po
c. ho.vo.a ‘get full’ hov.vo.a
d. ma.ve.ta ‘receive’ mav.ve.ta
e. yep.sa ‘arrive’ yeep.sa

The phonological expression of habitual in these cases is related to the marked heavy syllable reduplication, and involves the enhancement of a phonological string with an additional mora, realized via the “emergence of the unmarked”. In Yaqui, the usual way to express an additional mora is through gemination. However, as in example (19e), when there is a medial consonant cluster gemination cannot occur, presumably for phonotactic reasons (e.g. an illicit consonant cluster would result with gemination: *yeppsa), so the next best (i.e. the next least-marked) thing to do is lengthen the initial vowel.

4.2.2.2. Mora affixation in Tepecano (SUA)

Tepecano is an extinct language formerly belonging to the Tepiman sub-group.

According to Mason (1916), reduplication was used to mark the plural in the default case.

Some examples (without accents) are given in (20) (data from Mason 1916):

(20) Tepecano light syllable reduplication (= ‘plural’)

a. a.toc.kar ‘seat’ a.t.a.toc.kar
b. óc ‘corn-field’ óc
b. upp ‘skunk’ upp
d. nov ‘hand’ nov
e. du:r ‘ant’ du:du:r

However, Mason noted that “a second type of plural formation is found with disyllabic stems where the change. . .occurs within the stem itself” (p. 330). According to Mason,
this plural formation included both gemination and glottal-insertion, and at times included both processes simultaneously. I have not been able to discern an obvious semantic basis for the various classes.  

(21) Tepecano morphological gemination (= ‘plural’)
   a. i::pu:rr  ‘skirt’   i:p.purr
   b. ko:ko:n  ‘crow’   kok.kon
   c. hi.ku:rr  ‘peyote’   hik.kur
   d. ho:dai  ‘stone’   hod.dai
   e. ia.puc.kar  ‘sweat-cloth’   iap.puc.kar  

(22) Tepecano glottal-insertion (= ‘plural’)
   a. ta.tak  ‘nerve’   ta’t.tak
   b. go.goc  ‘dog’   go’goc
   c. i.ma.i  ‘squash’   i’ma.i
   d. du.dur  ‘jaguar’   du’dur
   e. a.s.a.k  ‘net’   a’sak

In Mason’s data there are two homophonous forms, one with reduplication and the other without: dudu:r ‘ant (pl.)’ and dudu:r ‘jaguar (sg.).’

Mason discusses the difficulty that he had in actually hearing the expression of some of the manifestations of mora affixation, and notes that the transcriptions that he used when recording these examples might not be accurate:

   [w]hile it is a delicate task to analyze correctly the phonetic characteristics of these plurals, and some of the above are probably not accurately expressed, yet the general process of secondary plural formation seems to be. . . that a medial stop is lengthened or a glottal stop introduced. . .” (p. 331, emphasis added).

---

13 Hill and Zepeda (1994) have identified a semantic class for marked plurals in another Tepiman language: Tohono O’odham (see discussion in section 4.4). This leads one to suspect that there might have also been one in Tepecano, as well as in Southeastern Tepehuan (cf. Kager 1999).
Haugen (in press a) draws two conclusions from the Tepecano data. First, there must have been a process of morphological gemination in this language. The phonetic realization seems to have been ambiguous between consonant doubling and laryngeal-insertion. Although it might seem unlikely that it had both processes, whether it did or not will probably never be known for sure.

Second, like its sister Tepiman languages Tepecano had a complex system of marking plural nouns, but whereas extant Tepiman languages use a distinction between short and long vowels, Tepecano used short vowel reduplication and mora affixation. With the loss of this language the historical development of reduplicative allomorphy (not to mention semantic classes) in Tepiman, and Uto-Aztecan more generally, has been obscured.

4.2.2.3. Mora affixation in the Numic languages (NUA)

McLaughlin (2001) provides evidence that the Numic languages also have the process of mora affixation. The morphological effects of aspectual gemination are independent of phonological gemination brought about by the “final features” of certain Numic stems. Elzinga (1999) claims that these geminates are non-moraic in Goshute Shoshone, since they do not affect (and are not affected by) the assignment of stress. However, I assume that the morphological gemination is moraic—the gemination is the morpheme; i.e. the morpheme is a bare mora affix.\textsuperscript{14}

\textsuperscript{14} This is similar to my earlier claim that coda consonants are also not moraic in Yaqui, unless they are there for a morphological purpose. The comparative aspects of syllable weight in UA would be an interesting topic for further investigation, and may have some import for historical reconstruction (e.g. impacting certain claims about PUA stress patterns, such as those of Manaster Ramer 1993), but this goes beyond the scope of our investigation here.
In Numic, mora affixation marks various aspectual distinctions as well as agreement. McLaughlin reconstructs the proto-Numic pattern to be gemination indicating momentaneous aspect, singular (ergative) agreement, and “move while doing” as opposed to “move to do”. This morphological gemination either appears on the verb stem or on motion suffixes. Examples from Central Numic are given in (23):

(21) Central Numic (McLaughlin 2001)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panamint</td>
<td>a.</td>
<td>kati</td>
<td>‘sit’</td>
<td>kati</td>
</tr>
<tr>
<td></td>
<td>b.</td>
<td>yake</td>
<td>‘cry’</td>
<td>yak ke</td>
</tr>
<tr>
<td>Shoshone</td>
<td>c.</td>
<td>kati</td>
<td>‘sit’</td>
<td>kati</td>
</tr>
<tr>
<td></td>
<td>d.</td>
<td>yakai</td>
<td>‘cry’</td>
<td>yak.kai</td>
</tr>
<tr>
<td>Comanche</td>
<td>e.</td>
<td>kati</td>
<td>‘sit’</td>
<td>kati</td>
</tr>
</tbody>
</table>

These examples from Central Numic illustrate the instantiation of mora affixation as morphological gemination.

These data from Central Numic lead to an interesting problem, however. The mora affixation analysis, which would imply that the first syllable is made heavy after affixation, leads to the expectation that stress will be attracted to that first syllable. Thus, under standard analyses of quantity-sensitive stress, a stem that normally has second syllable stress should surface with first syllable stress if that first syllable becomes heavy: e.g. hypothetical kati → kāt. ti, because of the heaviness (i.e. moraic quantity) of the first syllable. However, exactly the opposite seems to occur with durative gemination in the Goshute dialect of Shoshone (Dirk Elzinga, personal communication):
(22) **Unexpected stress shift in Goshute durative gemination** (Dirk Elzinga, p.c.)

- a. hipí ‘drink’ → hippí ‘drinking’
- b. káte ‘sit’ → katté ‘sitting’
- c. yákaí ‘cry’ → yakkái ‘crying’
- d. wéne ‘stand’ → wenne ‘standing’
- e. hápi ‘lie’ → happí ‘lying’

As noted above, Elzinga argues that Goshute coda consonants are not moraic since they do not attract stress. Under the mora affixation analysis that I have proposed for Uto-Aztecan, where the coda consonant crucially is moraic, the shifting of stress in these cases is quite mysterious.

One solution to the problem would be to simply say that Goshute is a quantity-insensitive language. Another possibility has been suggested by Hayes (1995: 299-305), who employs a layered representation of moraic structure, similar to other proposals involving stress grids. The purpose of Hayes’s novel representation is to capture languages that make dual distinctions of weight, such as CV: syllables that count as heavy but with CVC syllables that do not, or CVC syllables that count as heavy for some processes and as light for others.\(^{15}\)

Regardless of the correct synchronic analysis of the interaction of weight and stress in Central Numic, I think that it is clear that the forms discussed here show indisputable evidence of at least a diachronic process of mora affixation, reflected in Central Numic with morphological gemination. In Western Numic we see different reflexes of this mora affixation.

\(^{15}\) I thank Dirk Elzinga for pointing me to the Hayes (1995) analysis.
McLaughlin (2001) gives Northern Paiute examples showing the devoicing of a segment to indicate ‘durative’. The appearance of a voiceless stop in intervocalic position here indicates underlying gemination, since we would otherwise expect regular spirantization (i.e. fricativization) in this position (Snapp et al. 1982):

(23) Western Numic

Northern Paiute (McLaughlin 2001)

a. yaga ‘cry’ → yaka ‘cry (dur.)’
b. zoba ‘gather’ → zopa ‘gather (dur.)’

Snapp et al. (1982) also include the following for ‘durative’ in Northern Paiute:

(24) Other examples of ‘durative’ in Northern Paiute (Snapp et al. 1982)

a. mia → mi’a ‘go’
b. nimi → nimmi ‘move about/walk/go’
c. himma → hi’ma ‘carry (pl.)’
d. sumaya → suma’ya ‘remember’

Thus, in Northern Paiute, in addition to gemination, we can also see epenthesis of glottal stop and epenthesis of a copy of the final vowel of the word for the durative morpheme.

These reflexes of mora-insertion are independent from the syllabic reduplication used for number agreement, which is prefixal for ‘dual’ (e.g. mia → mi.mia ‘go (dual)’), and, occasionally, suffixal for ‘plural’ (e.g. mia → mia.a ‘go (pl.)’) (Snapp et al. 1982).

Once again, in all of these cases we must consider the prosodic status of the morphemes under consideration in order to make the claim that the same process, i.e. mora affixation, is occurring in all of these languages, even though the exact phonological realization of the additional mora may be language-specific. In the next section I will argue that the best way of capturing the language-specific details is by the language-specific relative ranking of universal constraints.
4.3. Capturing the Uto-Aztecan generalizations through OT constraint-ranking

The advantage that Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993, 1995) has over earlier, arbitrary rule-based models is its overt focus on capturing cross-linguistic generalizations. By positing universal constraints on phonological form that can be differentially ranked relative to one another in particular grammars, we are also able to describe language-specific peculiarities. Thus, examining closely-related languages provides us with a way to clearly examine the system of markedness (i.e. well-formedness) constraints and their interaction with other constraints, such as faithfulness to underlying form or to the base of reduplication.

The divergent patterns of mora augmentation in most of the Uto-Aztecan languages can be easily explained by an appeal to the ranking of three constraints, given in (25), where the surfacing reduplication pattern is a result of the "emergence of the unmarked" (TETU) (McCarthy and Prince 1994).

(25) a. DEP-IO: No epenthesis.
   b. *LONG-V: No long vowels.
   c. *LONG-C: No gemination.

With just these three constraints we can (almost) account for the full typology of heavy syllable reduplication found in Uto-Aztecan (see Table 4.1). This typology accounts for those cases of heavy syllable reduplication which only copy from the first syllable, and not from the second syllable (e.g. in the Mayo accented class):
Table 4.1. Markedness typology for Uto-Aztecan heavy syllable reduplication

<table>
<thead>
<tr>
<th>Constraint Ranking</th>
<th>Result</th>
<th>Example Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. DEP-IO, *LONG-V &gt;&gt; *LONG-C</td>
<td>gemination</td>
<td>Yaqui, Central Numic, Tohono O’odham</td>
</tr>
<tr>
<td>*b. DEP-IO, *LONG-C, &gt;&gt; *LONG-V</td>
<td>vowel-lengthening</td>
<td>Nahuatl, Tohono O’odham</td>
</tr>
<tr>
<td>c. *LONG-V, *LONG-C &gt;&gt; DEP-IO</td>
<td>epenthesis ([ʔ] or [h])</td>
<td>Guarijio, Nahuatl, Mono</td>
</tr>
</tbody>
</table>

Because long-vowel reduplication only seems to appear in languages that also have one of the other forms of heavy syllable reduplication, the ranking illustrated in (b) in Table 4.1 might actually be unattested. For example, in Tohono O’odham the unmarked heavy syllable reduplication is of the type which induces gemination, and long vowel reduplication only occurs with a marginal class of a few forms with a specific, culturally-relevant semantic motivation. There might be a historical explanation for this unattested ranking, if it is indeed unattested, having to do with a possible phonemic vowel-length distinction in PUA (as reconstructed, for example, by Langacker 1977a)—a vowel length distinction that is maintained in some daughter languages.16

Based on the wide distribution of the gminating reduplication across the UA languages, and following the principle of economy in historical reconstruction, I propose that PUA had the constraint-ranking given in Table 4.1(a), and that the default heavy syllable reduplication and mora affixation patterns involved gemination. Some languages have independently undergone subsequent constraint re-rankings. The most common re-

16 Phonemic vowel length in PUA is not universally accepted. See Manaster Ramer (1993) for a stress-based account of long vowels in some Uto-Aztecan languages.
ranking is that of Table 4.1 (e), forcing epenthesis of glottal stop or [h] (e.g. in Guaricjio and Nahuatl).

The following tableaux illustrate the interaction of the three constraints under consideration here. In each of these simple cases we only need to specify the one constraint that is lowest-ranked. This is the constraint that determines the winning candidate. The relative unimportance of the ranking between the two higher constraints is illustrated in the contrast between the (a) and (b) pairs in the following examples—the winning candidate is chosen regardless of which constraint is highest-ranked. Thus, only the lowest-ranked constraint is relevant here. However, based on other considerations in each of the languages involved, discussed below, I have indicated a plausible preference for constraint-ranking among the two higher constraints by labeling the relevant tableau with a "(26) Geminating Reduplication: Yaqui.

i.

<table>
<thead>
<tr>
<th>/RED2 + bwiika/</th>
<th>*LONG-V</th>
<th>DEP-IO</th>
<th>*LONG-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.  ☐ bwib.bwi.ka</td>
<td>*LONG-V</td>
<td>DEP-IO</td>
<td>*LONG-C</td>
</tr>
<tr>
<td>b.  bwii.bwi.ka</td>
<td>*!</td>
<td>DEP-IO</td>
<td>*LONG-C</td>
</tr>
<tr>
<td>c.  bwi? bwi.ka</td>
<td>*!</td>
<td>DEP-IO</td>
<td>*LONG-C</td>
</tr>
</tbody>
</table>

ii. ☐ Phonemic long vowels

<table>
<thead>
<tr>
<th>/RED2 + bwiika/</th>
<th>DEP-IO</th>
<th>*LONG-V</th>
<th>*LONG-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.  ☐ bwib.bwi.ka</td>
<td>DEP-IO</td>
<td>*LONG-V</td>
<td>*LONG-C</td>
</tr>
<tr>
<td>b.  bwii.bwi.ka</td>
<td>DEP-IO</td>
<td>*LONG-V</td>
<td>*LONG-C</td>
</tr>
<tr>
<td>c.  bwi? bwi.ka</td>
<td>DEP-IO</td>
<td>*LONG-V</td>
<td>*LONG-C</td>
</tr>
</tbody>
</table>

17 Symbols used in tableaux are as follows: ☐ = a candidate correctly chosen by a given constraint-ranking; ☐ = an actual surfacing output form that is not selected by a given constraint-ranking; and ☐ = a candidate incorrectly chosen by a given constraint-ranking.
(27) Glottal-epenthesis Reduplication: Guarijío

i. ☺ No geminates, long V's can occur

<table>
<thead>
<tr>
<th>/RED + naó/</th>
<th>*LONG-C</th>
<th>*LONG-V</th>
<th>DEP-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nan.náo</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. naa.náo</td>
<td></td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>c. naʔ.náo</td>
<td></td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

ii.

<table>
<thead>
<tr>
<th>/RED + naó/</th>
<th>*LONG-V</th>
<th>*LONG-C</th>
<th>DEP-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nan.náo</td>
<td></td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>b. naa.náo</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. naʔ.náo</td>
<td></td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

As indicated above, facts specific to each of these languages could lead us to decide on the relative ranking of the two higher constraints. For example, Yaqui has phonemically long vowels, so perhaps DEP-IO is ranked above *LONG-V (thus “☺” for 26-ii). In Guarijío, there are absolutely no geminates, but long vowels do occur. Therefore, it might be the case that the highest-ranked of these constraints is *LONG-C. What is relevant for the determination of the winning candidate here, however, is that in these cases the only constraint that decides on the output form of the reduplicant is the lowest-ranked constraint, since that alone is enough to decide what will surface in marked heavy syllable reduplication.

One of the benefits of an OT-analysis is that the constraints proposed above apply to the language as a whole, rather than being reduplication-specific. Thus, the constraint-rankings proposed here should have implications for the rest of the phonology of these languages. This is indeed borne out. For example, at some point in its history Guarijío acquired a general ban on geminates, which can be represented as the promotion of a single constraint: *LONG-C. For Guarijío this is independently needed, as the
correspondence between Guarijio [h] and Mayo (and Yaqui) gemination attests (Manaster Ramer 1992, Dakin 1996):

(28) | Guarijio | Mayo |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ehté</td>
<td>étte</td>
<td>'louse'</td>
</tr>
<tr>
<td>b. ehka-ní</td>
<td>hékka</td>
<td>'shade'</td>
</tr>
<tr>
<td>c. kahti</td>
<td>káttek</td>
<td>'to be seated'</td>
</tr>
<tr>
<td>d. mahtá</td>
<td>máttta</td>
<td>'metate'</td>
</tr>
<tr>
<td>e. tehté</td>
<td>tétte</td>
<td>'stone'</td>
</tr>
</tbody>
</table>

By looking at only these two languages it is impossible to tell which is the innovator with respect to these two constraints, but in the larger context of the family, it seems likely that that most languages probably preserved the original ranking—i.e. the Yaqui/Mayo ranking.\(^\text{18}\)

One cautionary note should be made regarding the analysis of these patterns from the OT framework. This is the fact that we cannot assume that two languages that share the same ranking of any particular constraints, e.g. Guarijio and Mono with the constraints under consideration here, shared in a single innovating constraint-re-ranking event. That is, sharing a particular ranking between two constraints cannot be taken as evidence for family tree sub-branching, for example. Constraint-ranking is evidence for relationships only in the context of other evidence—very closely-related languages (e.g. Yaqui and Mayo, which were regarded as a single language, “Cáhita”, in the colonial era) can be

\(^{18}\) It is also possible to re-frame this as the development of geminates in Mayo, rather than the development of a ban on geminates in Guarijio, wherein Mayo has demoted the constraint *LONG-C rather than Guarijio having promoted it. Since I assume that PUA had geminating heavy-syllable reduplication, however, it must have been the other way around if there have not been intervening constraint re-rankings between PUA and modern Guarijio and Mayo, of which we have no positive evidence. I should also point out that it has been suggested that phonological change always occurs through constraint-promotion (Kennedy 2001).
relatively dissimilar in their surface phonology, derived by underlying constraint-rankings, and still be (relatively) mutually intelligible.

The OT constraint-ranking approach makes a strong prediction. Since these constraints are general and not specific to the process of reduplication or mora affixation, it is predicted that, all else being equal, languages that have both marked heavy syllable reduplication and bare mora affixation should realize the extra mora in the same way—gemination, vowel lengthening, or epenthesis. Again, this is generally borne out.

(29) Geminating mora affixation: Yaqui

i.

<table>
<thead>
<tr>
<th>/μ + maveta /</th>
<th>*LONG-V</th>
<th>DEP-IO</th>
<th>*LONG-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. mav.veta</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. maa.veta</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ma?.veta</td>
<td>!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ii. Phonemic long vowels

<table>
<thead>
<tr>
<th>/μ + maveta /</th>
<th>DEP-IO</th>
<th>*LONG-V</th>
<th>*LONG-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. mav.veta</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. maa.veta</td>
<td></td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>c. ma?.veta</td>
<td></td>
<td>!</td>
<td></td>
</tr>
</tbody>
</table>

(30) Glottal-epenthesis Mora Affixation: Tepecano (assuming glottal-insertion)

i.

<table>
<thead>
<tr>
<th>/μ + dudur /</th>
<th>*LONG-C</th>
<th>*LONG-V</th>
<th>DEP-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dud.dur</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. duu.dur</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. du?.dur</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

ii.

<table>
<thead>
<tr>
<th>/μ + dudur /</th>
<th>*LONG-V</th>
<th>*LONG-C</th>
<th>DEP-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dud.dur</td>
<td></td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>b. duu.dur</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. du?.dur</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The approach advocated here predicts that, all else being equal, heavy syllable reduplication and mora affixation will trigger the same process: gemination, vowel-lengthening, or epenthesis of an unmarked consonant. However, this might be violated if phonological contexts differ, as in Yaqui words with word-medial consonant clusters receiving vowel-lengthening (e.g. *yepsa* → *yeeepsa* ‘arrive’), or possibly also in areas with a moribund language, as is possibly the case with Tepecano. Although a rule-based account could be invoked to account for these data, it is not clear how such an approach would account for the systematic variation that we have seen.

An additional advantage to this system is that the same constraint-ranking should apply to all aspects of the phonology of any individual language.

4.4. Some complicating data: A three-way distinction in Tohono O’odham

One Uto-Aztecan language that muddies the waters of the account hitherto sketched is Tohono O’odham (formerly Papago), a Tepiman language. This language has multiple reduplicative duplemes, including two different patterns of marked heavy syllable reduplication. One heavy syllable reduplication pattern surfaces with a long vowel for a special class of human plural nouns, and the other marks the heavy syllable with gemination to indicate other meanings: distributive, repetitive, etc.

According to Hill and Zepeda (1994), the usual pattern of plural-marking in Tohono O’odham is with a light CV- syllable reduplicative prefix:19

---

19 There is some debate as to whether or not the reduplicant is a prefix or an infix in these cases. Riggle (2001) has argued for an infixation analysis for the closely-related language Pima (Akimel O’odham). Other scholars (e.g. Hill and Zepeda 1994, Fitzgerald 2000) argue for an analysis with prefixation and syncope in the base for Tohono O’odham. There are several reasons to accept a prefixal analysis, not the
(31) Tohono O’odham plurals (Hill and Zepeda 1994)

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bitokoi</td>
<td>bibtokoi</td>
<td>‘Pinacate beetle’</td>
</tr>
<tr>
<td>b. ce:kol</td>
<td>ceckol</td>
<td>‘squirrel’</td>
</tr>
<tr>
<td>c. cu:wí</td>
<td>cu:cu:wí</td>
<td>‘jackrabbit’</td>
</tr>
<tr>
<td>d. gogs</td>
<td>gogogs</td>
<td>‘dog’</td>
</tr>
<tr>
<td>e. mi:stol</td>
<td>mimstol</td>
<td>‘cat’</td>
</tr>
</tbody>
</table>

With a special culturally-significant class of (less than 100) roots, however, the plural is marked with a long vowel:

(32) Tohono O’odham long-vowel plurals (Hill and Zepeda 1994)

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ban</td>
<td>ba:ban</td>
<td>‘coyote’</td>
</tr>
<tr>
<td>b. ma:d</td>
<td>ma:ma:d</td>
<td>‘woman’s child, younger sister’s child’</td>
</tr>
<tr>
<td>c. ñem</td>
<td>ñe:ñem</td>
<td>‘liver’</td>
</tr>
<tr>
<td>d. bahi</td>
<td>ba:bhai</td>
<td>‘tail’</td>
</tr>
<tr>
<td>e. şon</td>
<td>şo:şon</td>
<td>‘trunk of a plant’</td>
</tr>
</tbody>
</table>

In addition to this limited pattern of marked heavy syllable reduplication for plural nouns, there are also highly productive cases of marked heavy syllables that trigger gemination, as with distributive nouns and verbs (33) and repetitive verbs (34).

(33) O’odham distributives (with gemination) (Fitzgerald 2003)

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>distributive</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nowiu</td>
<td>nonowiu</td>
<td>nonnowiu</td>
<td>‘ox’</td>
</tr>
<tr>
<td>b. nahagio</td>
<td>nanhagio</td>
<td>nanhagio</td>
<td>‘earring’</td>
</tr>
<tr>
<td>c. hódai</td>
<td>hohodai</td>
<td>hohhodai</td>
<td>‘rock, stone’</td>
</tr>
<tr>
<td>d. ?a:g</td>
<td>?a?ag</td>
<td>?a??ag</td>
<td>‘a pair of animal horns’</td>
</tr>
</tbody>
</table>

(34) O’odham repetitive verbs (with gemination) (Fitzgerald 2003)

<table>
<thead>
<tr>
<th>unitative</th>
<th>repetitive</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. giw</td>
<td>giggiw</td>
<td>‘hit something’</td>
</tr>
<tr>
<td>b. hihim</td>
<td>hihhim</td>
<td>‘laugh’</td>
</tr>
<tr>
<td>c. huholgat</td>
<td>huholgat</td>
<td>‘menstruate’</td>
</tr>
<tr>
<td>d. kow</td>
<td>kokkow</td>
<td>‘dig something out of the ground’</td>
</tr>
<tr>
<td>e. jia</td>
<td>jinna</td>
<td>‘look in a certain direction’</td>
</tr>
</tbody>
</table>

Least of which is the fact that syncope is a quite general process in O’odham, and most all other instances of reduplication across Uto-Aztecan are prefixal.
I claim that the long-vowel reduplication in Tohono O'odham, which is only used in a small class of plural nouns, is a secondary development. The use of a long-vowel reduplicant is a way to “enhance contrast” between reduplicative morphemes (Urbanczyk 2002), in that nouns with gemination are typically construed as being ‘distributed’, whereas nouns with long vowels are construed as being ‘plural’.

Data like these from Tohono O’odham do bring up an interesting issue within Optimality Theory, though, since, as was mentioned before, the constraint-ranking should be general to the language as a whole, as opposed to being process- or morpheme-specific. Thus, we are left with a puzzle as to how to induce a long vowel reduplicant in a constraint hierarchy that usually induces gemination. I see two possible solutions to this problem.

The first is that this language, with both geminating and long-vowel reduplication, can be accounted for with the addition of a fourth (highly-ranked) constraint favoring a distinction between morphemes, as in (35):

\[(35) \text{ DISTINCTMORPHEME: No homophonous morphemes.}\]

The purpose of this constraint would be to cause vowel-lengthening in those cases where a plural noun would be confused with a distributive noun, in capitulation to a formally-encoded preference for homophony avoidance.

This approach is advocated by Czaykowska-Higgins and Urbanczyk (2001), who introduce the constraint DISTINCTFORM “to formally express the morphological goal that distinct meanings should have distinct forms” (p. 114). Under Czaykowska-Higgins and
Urbanczyk’s analysis, this constraint compares stems (reduplicants plus their bases and any other affixes) and evaluates them based on their phonetic similarity and semantic meanings. The constraint acts as a function that does a pair-wise comparison of stems, calculating the phonetic differences between them. If two identical phonetic strings have different meanings, then DISTINCTFORM is violated. Ironically, the long vowel vs. geminating heavy syllable reduplicants in O’odham seem to be stronger evidence in favor of this constraint than the data from Nxa’amxcin (Salish) that Czaykowska-Higgins and Urbanczyk themselves give, since the different Nxa’amxcin reduplicants that appear for the Augmentative (AUG) and Characteristic (CHAR) functions are prefixal and suffixal, respectively. I illustrate the workings of this constraint, using O’odham examples, in (36). (In this tableau, DIST = ‘distributive’ and M.PL. = ‘marked plural’; the underlined segments indicate the contrasting segments being identified by DISTINCTFORM in the second column):

(36) Evaluating DISTINCTFORM (~ Czaykowska-Higgins and Urbanczyk 2001: 116 [16])

<table>
<thead>
<tr>
<th>DIST + ROOT ↔ M.PL. + ROOT</th>
<th>DISTINCTFORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. © bab-ban ↔ baab-ban</td>
<td>{ b ↔ a }</td>
</tr>
<tr>
<td>b. 0 baa-ban ↔ bab-ban</td>
<td>{ a ↔ b }</td>
</tr>
<tr>
<td>c. bab-ban ↔ bab-ban</td>
<td>Ø</td>
</tr>
<tr>
<td>d. baa-ban ↔ baa-ban</td>
<td>Ø</td>
</tr>
</tbody>
</table>

Candidates (36)c and (36)d have identical output forms, so DISTINCTFORM turns out an empty set and is violated. Since candidates (36)a and (36)b each have a different phonetic realization they both satisfy DISTINCTFORM, but they do so equally and must be differentiated by other constraints.

In the Nxa’amxcin examples discussed by Czaykowska-Higgins and Urbanczyk, the reduplicants are distinguished by the fact that AUG is prefixal while CHAR is suffixal,
and that the former always surfaces with a reduced vowel (schwa) while the latter
characteristically has a vowel that is faithful to that of the base (i.e. the root to which it
attached):\textsuperscript{20}

(37) \textit{Nxa’amxcín Reduplication} (Czaykowska-Higgins and Urbanczyk 2001)

<table>
<thead>
<tr>
<th>Form</th>
<th>Semantics</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. AUG</td>
<td>q’ol-q’il-t</td>
<td>{AUG-, (\sqrt{\text{‘ill’}}), -stative} ‘they’re all sick’</td>
</tr>
<tr>
<td>b. CHAR</td>
<td>q’il-q’il-t</td>
<td>{(\sqrt{\text{‘ill’}}), -CHAR, -stative} ‘it hurts bad’</td>
</tr>
</tbody>
</table>

A constraint against prominent prefixes is able to derive the vowel-reduction in the AUG
reduplicant.

In the Tohono O’odham cases, however, the reduplicants appear in exactly the same
location within the word, prefixally, and both the geminating and vowel-lengthening
reduplicants are equally faithful to the base, in that neither mechanism used to create the
heavy syllable in these reduplicants satisfies Base-Reduplicant faithfulness (MAX-BR)
any better than the other (i.e. neither results in a more “full” copy of the base, as with the
unattested \(*\text{ban-han}\)). The constraint \textsc{DistinctForm}, as formulated by Czaykowska-
Higgins and Urbanczyk, does not result in the assignment of the reduplicating stem to the
correct morphological class; that is, this constraint is equally satisfied by two wrong
candidates, with the distributive output receiving a long vowel and the marked plural
receiving germination (as in 36b), as by two correct ones (cf. 36a).

In addition, the type of constraint evaluation performed by \textsc{DistinctForm} pre-
supposes that the phonological component of the grammar has access to other output

\textsuperscript{20} The notation in (37) is slightly different than that used by Czaykowska-Higgins and Urbanczyk; the
reduplicant here appears in bold.
candidates in a given language. However, this brings up troublesome theoretical questions about what it is that constraints can evaluate, and whether constraints can compare output candidates in one tableau to candidates in other tableaux. Further, this approach is applicable to a process-based and word-and-paradigm models of grammar, and seems to be inconsistent with the piece-based approach adopted here (cf. Urbanczyk 2002).

In sum, the actual morphological classes that the various O'odham reduplicants belong to must be indicated in some way independently of DISTINCTFORM, therefore this constraint does not do any actual work for the analysis of reduplication, at least in this language. It merely stipulates that semantically different morphemes must be phonetically different, which already has to be stipulated somewhere else in the grammar, in order to derive long vowels instead of gemination.

Another possibility is to make constraint evaluation morpheme-specific (i.e. "lexical"), in a fashion similar to what is proposed in Direct Optimality Theory (DOT) (Golston 1996). In DOT, the lexical representation of each morpheme is a set of constraint-violations. A more advantageous approach would be to lexically flag an entire morphemic class for a special phonological treatment. Archangeli (2003), for example, accounts for irregularities in the expression of vowels in some words in Tiv (Niger-Congo) by lexically encoding necessary violations of highly-ranked constraints. The second vowel in Tiv CVCV words typically agree for height and rounding with the first vowel, as long as that vowel is not low ([a] or [ɔ]), in which case the second vowel appears as a mid vowel ([e] or [o]). According to Archangeli, "the robust class of
exceptions, in which the initial vowel is high \([i.e. \text{[i]} \text{ or } \text{[u]}—\text{JH]}\) and the final vowel is low or mid, is used to motivate the rightward alignment of \([+\text{round}]\) to \([+\text{high}]\), preferred over the more general rightward alignment of \([+\text{round}]\)” (p. 234). Following the DOT notion of lexically requiring a constraint violation, in this case *Align[+Hi]-Right, candidates such as *imbise, without a high vowel at the end of the word (i.e. without the feature \([+\text{high}]\) aligned to the right edge of the word), are preferred over candidates such as *[imbisi], expected with vowel harmony, and *[imbese], with two violations of alignment of the feature \([+\text{high}]\). Archangeli thus formally encodes the quirky realization of vowels that would otherwise succumb to processes of vowel harmony.

This approach could be extended to the O’odham marked plurals by insisting, through a lexical stipulation (independently necessary since there are sociocultural reasons for nominal roots to fit into this class), that the plurals of these forms violate the constraint *LONG-V. An illustrative tableau is given in (38):

(38)

<table>
<thead>
<tr>
<th></th>
<th>RED=σ_{hi}</th>
<th>DEP-IO</th>
<th>*LONG-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ba-ban</td>
<td>*!</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>b. bab-ban</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>c. baa-ban</td>
<td></td>
<td>*L</td>
<td></td>
</tr>
<tr>
<td>d. ba?-ban</td>
<td></td>
<td>*!</td>
<td>*!</td>
</tr>
<tr>
<td>e. baa-baan</td>
<td></td>
<td></td>
<td><em>L</em>!</td>
</tr>
</tbody>
</table>

In the underlying representation in (38), the root *ban ‘coyote’ is flagged for membership in the nominal category that takes long vowel reduplicants, indicated with a subscript [MP], for ‘marked plural class’. This membership is formally encoded by the lexically-specified violation of *LONG-V, also subscripted with [MP], which also appears in the underlying representation. This constraint must be violated, and since this
violation is required it does not count against a candidate. Candidates that do not have a violation, as with (38)a, b and d, all incur inverse violations (*₁), which would rule them out if they were not already ruled out on independent grounds. Candidate (38)a violates the requirement that the reduplicant appear in a heavy syllable (RED=σ₁), and candidate (38)c has a fatal violation of DEP-IO with its epenthetic glottal stop. The lexically-specified violation of *LONG-V does correctly rule out the candidate that we would otherwise expect to be optimal, (38)b, which is what surfaces in other contexts (e.g. in the distributive or repetitive forms). Candidate (38)c has one violation of *LONG-V too many, and it is therefore less optimal than candidate (38)c.

Although this account is stipulative, the stipulation is only made for the class of lexically marked items that we already have to flag as exceptional. This stipulation rules out both reduplicants with short vowels, as appear with unmarked plurals in O'odham, as well as heavy syllable reduplicants with gemination, as appear in the distributive. Reduplicants with epenthesis are ruled out by a high-ranked DEP-IO, which is also independently necessary (cf. the examples of marked heavy syllable reduplication in Guarijío in section 4.2.1.3).

The crucial thing here is that heavy syllable reduplication involves the reduplication of heavy syllables, at least in Uto-Aztecan, regardless of the surfacing phonological realization of those syllables. Tohono O’odham long vowel reduplicants are simply a variation on geminating reduplicants utilized to mark a different meaning (i.e. dupleme). In this analysis I agree with Urbanczyk (2002), who proposes that the morphemes that are morpho-phonologically enhanced are those that are innovative.
Although data like those in O'odham call for a theory-internal solution, the general picture of constraint-ranking is supported by these data. In O'odham we see vowel-lengthening rather than epenthesis in this kind of reduplication, which fits with the general picture of constraints on word-size in O'odham (Fitzgerald 1997). Specific aspects of Optimality Theory could be abandoned or supplemented. For example, one could debate the necessity of the idea of a single Input-Output relation, as opposed to having a level-ordered grammar, as McCarthy and Prince (1993: 24) suppose might be necessary for prefixes and suffixes in Axininca Campa, and as is posited by Stonham (1994), without altering the over-all advantage of viewing phonology as the interaction of ranked constraints.

4.5. Against the alternative analysis

We began this chapter with a brief over-view of the Modular-Derivational theory of reduplication proposed by Raimy (2000), which crucially argues that there is no prosody and that reduplication is rule-based as opposed to constraint-based. I have tried to show that prosody is crucial to understanding the comparative aspects of reduplication in Uto-Aztecan. For example, without reference to a reduplicant composed of a heavy syllable, we could not reconstruct a 'distributive' morpheme of that form for PUA. However, I think that the evidence for this reconstruction is compelling, as I have tried to illustrate.

My argument against Raimy's theory is not that it cannot account for the kinds of reduplication that we see across Uto-Aztecan. In fact, the model proposed in Raimy (2000) is a very powerful (if not unconstrained) system, and his theory could account for
each of the patterns of data that we have encountered so far. I will briefly sketch the kind of analysis that his system would have to make for the comparative data that we have seen. I will use marked heavy syllable reduplication as the illustrative case. My conclusion will be that while a Raimy-style analysis is possible, the OT-based analysis is more perspicuous. This section is organized as follows. Section 4.5.1 demonstrates that prosody is crucial in the consideration of reduplication in Uto-Aztecan; section 4.5.2 considers the advantages of constraint-ranking when considering the comparative Uto-Aztecan data; and section 4.5.3 revisits the issue of the morphemic status of the reduplicant.

4.5.1. Reduplication and prosodic units

Raimy uses data from Tohono O’odham to argue for his model of reduplication, the crucial aspect of which is the explicit representation of linear order. For example, the underlying representation of an unreduplicated word like *pualt* ‘a door’ is as given in (38) (# = ‘beginning of string’ and % is ‘ending of string’):

\[(39) \# \rightarrow p \rightarrow u \rightarrow a \rightarrow t \rightarrow %\]

When the phonological rule of Linearization applies to this representation, the final output is *pualt*. In Raimy’s theory Linearization is a phonological rule that may be ordered with respect to other rules in a language, and in this way Raimy yields the familiar under- and over-application effects that we see cross-linguistically with reduplication.
What is especially of interest here is the way that reduplication structures are derived in this system. A given morpheme will have an underlying representation like that given in (39). When reduplication occurs, what happens is that another morpheme, e.g. ‘plural’, is combined with the morpheme in question, e.g. *pualt* ‘a door’. In traditional Optimality Theory it is generally assumed that the reduplicant itself is the morpheme, and this assumption goes back at least to the work of Marantz (1982) and Moravcsik (1978)—see section 4.5.3. However, in Raimy’s theory the “reduplicant” is actually an epiphenomenon of a re-adjustment rule triggered by a (usually) null affix. The “re-write” involved is the addition of a “jump link” (i.e. an additional arrow) into the underlying representation of the affected morpheme. This is shown in (40):

(40) \[ \# \rightarrow p \rightarrow u \rightarrow a \rightarrow l \rightarrow t \rightarrow \% \]

In this case, when the rule of Linearization goes so far as \( \# \rightarrow p \rightarrow u \), the additional link causes the repetition of the \( p \) and the \( u \). (The link is only followed once, thus infinite loops are ruled out by an independent principle of economy). The resulting output form, completely linearized, is *pupualt* ‘doors’. In Raimy’s system there is no reference to prosody. The rule that derives a CV reduplicant can be paraphrased as “the first vowel precedes the segment which precedes the first vowel”—see Raimy (2000: 113) for the formalization of this stipulation.

This is how the normal process of reduplication works within Raimy’s theoretical framework. Tohono O’odham is particularly interesting within Raimy’s theory, or any other theory for that matter, because of the additional complications of reduplication
within this language. First, in some instances we see syncope of a base vowel (41), and with the class of "heavy" (i.e. bimoraic) diphthongs (Miyashita 2002) we see the splitting of the diphthong (42) (data from Raimy 2000):

(41) Syncope in the base

a. babad 'frog' \(\rightarrow\) ba-bba\(\delta\) [bab.ba\(\delta\)]
b. gaso 'a fox' \(\rightarrow\) ga-gso [gag.so]
c. maskal 'a bandana' \(\rightarrow\) ma-mskal [mam.skal]

(42) Splitting of (heavy) diphthongs

a. doakag 'a life' \(\rightarrow\) do-dakag [do.da.kag] *[do-doa.kag]
b. koa 'forehead' \(\rightarrow\) ko-ka [ko.ca] *[ko-koa]
c. čioj 'boy, man' \(\rightarrow\) či- čož [či.čož] *[či.čioj]

In Raimy's theory, these forms are derived by adding jump links to the representation.

Considering here only the cases of reduplication with syncope, the final output form would be derived with the following series of operations:

(43) a. \# \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) \(\delta\) \(\rightarrow\) 

b. \# \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) \(\delta\) \(\rightarrow\) 

c. \# \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) \(\delta\) \(\rightarrow\) 

d. \# \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) b \(\rightarrow\) a \(\rightarrow\) \(\delta\) \(\rightarrow\)

(43)b and (43)c are stages of the phonological derivation where an affix adds jump links to the underlying representation. The ordering of the following of jump links in (43)c must be the \(a \rightarrow b\) jump link based on "the principle to spell out as many morphological links as possible" (p. 115). If the \(b \rightarrow b\) jump link (added in 43b) were followed first, then the \(a \rightarrow b\) jump link would not be followed at all; however, it should
be mentioned that there does not seem to be an independent way to preclude making the kinds of links which would be phonotactically impossible in a given language, such as the form \((bb\,\alpha\,\delta)\) that would occur if the other jump link were to be followed first (or if there were no \(a \rightarrow b\) jump link at all).

In other words, there is no principle limiting the kinds of jump links that can be created—they are arbitrary paths induced in the phonological representation. They have no connection to the attested patterns, supposedly epiphenomenal, of prosody which have been identified in the literature (e.g. in McCarthy and Prince 1986 and much subsequent work following the tradition of Prosodic Morphology, within or without OT itself). Thus, the fact that we see the same prosodic units showing up across Uto-Aztecan reduplicative allomorphy—light syllable, heavy syllable, disyllable, and mora affix (this last being incoherent in Raimy’s system, although he can and does posit an X-slot for “moraic” effects), is a complete accident.

One could easily imagine a Raimy-style theory which includes prosody, however; in fact, Frampton (2003) proposes such a theory. Therefore, we must also consider the other aspects of the Modular-Derivational theory with respect to the comparative Uto-Aztecan data.

4.5.2. Constraint-ranking, or no?

As mentioned briefly in the last section, Raimy can derive moraic effects by means of the insertion of a link to an X-slot within the underlying representation. He does this with several cases of cross-linguistic epenthetic elements in CVC reduplication. For example,
the related Austronesian languages Ilokano, Mokilese, and Kusaiean all have a heavy syllable prefixal reduplicant, each of which shows some variation in how it is realized (see Raimy 2000:127-139 for detailed discussion).

Since the variation in those Austronesian languages is somewhat (although not entirely) similar to the Uto-Aztecan marked heavy syllable reduplication patterns, I will extrapolate from Raimy’s Austronesian discussion to show how his theory would account for the Uto-Aztecan patterns. The crucial difference between the Austronesian and Uto-Aztecan cases is one of the largest weaknesses of Raimy’s theory, since while Austronesian languages typically copy into the second syllable to make a heavy syllable prefix (e.g. Ilokano pu.sa ‘cat’ → pus-pu.sa ‘cats’), Uto-Aztecan languages typically do not. Raimy’s theory has no non-arbitrary way to capture this generalization, since the default assumption of “the-full-stem-is-the-base-for-reduplication” is apparently not maintained in Uto-Aztecan.

Nevertheless, arbitrary rules could be assigned to derive the Uto-Aztecan marked heavy syllable reduplication patterns.

For the cases of heavy syllable reduplication with gemination and long vowels, an additional “self-loop” readjustment rule could be added to the reduplicant. The use of such a self-loop, the standard way of deriving gemination in Raimy’s system, are shown for each case in (44) and (45). (44) is Raimy’s analysis of the data (figure 110c, p. 119); (45) is an extrapolation based on Raimy’s analysis in (44).
(44) **Tohono O’odham Long Vowel Plural Reduplication**

\[ \text{ban} \quad \text{‘coyote’} \]

\[
\begin{array}{c}
\# 
\rightarrow b
\rightarrow a
\rightarrow n
\rightarrow \%
\end{array}
\rightarrow \text{baar-ban}
\]

(45) **Tohono O’odham Geminating Distributive Reduplication**

\[ \text{nowiu} \quad \text{‘ox’} \]

\[
\begin{array}{c}
\# 
\rightarrow n
\rightarrow o
\rightarrow w
\rightarrow i
\rightarrow u
\rightarrow \%
\end{array}
\rightarrow \text{non-nowiu}
\]

For (45), Raimy proposes an economy principle for choosing among the different possible orders for following the two different links. According to Raimy,

Linearization of [45] shows that the added self-loop is followed before the back link is because it results in a linearization that discharges the added links quicker than other possible linearizations. If the back link were followed first then a form with a long vowel in the stem would result. This is a less economical linearization than the occurring surface form because it takes traveling five precedence links to discharge all added links as opposed to the occurring linearization that only requires four precedence links to discharge these links.

(pp. 119-120)

I illustrate the contrasting representations in (46), where the added links appear doubled and in bold:

(46)  

a. \[ \# \rightarrow b \rightarrow a \rightarrow b \rightarrow \ldots \]

b. \[ \# \rightarrow b \rightarrow a \rightarrow b \rightarrow a \rightarrow \ldots \]
Raimy’s use of this “economy principle” to encode “quickness” of discharging links seems to be completely arbitrary. For example, why is the discharge of all the links preferred over other possible linearizations, such as “get to the end (%) as quickly as possible”, in which case simply ignoring any added links would always be more “economical” than following them? There are any number of conceivable alternative notions of “economy” that could be applied in these situations, and it is not clear how to non-arbitrarily choose the correct one.

In any event, this principle does not have any effect at all on those cases of reduplication with gemination, since the number of links necessary to “discharge the added links” is identical regardless of which loop is followed first, the from-the-first-vowel-back-to-the-consonant-before-the-first-vowel link, or the self-link on the first consonant:

(47) a. # → n → o → n → n → . . .
b. # → n → n → o → w → i → . . .

An independent rule referring to the phonotactics of the language is necessary to rule out (47)b.

Raimy’s approach also does not associate these two reduplicative processes, which from a prosodic standpoint we can clearly identify as being two different instantiations of heavy syllable reduplication. Since Raimy’s theory proposes these linkings as arbitrary aspects of the stem involved, it is difficult to image how he would capture the possibility of a stem that could reduplicate with both a long vowel and with gemination, if such a stem indeed exists.
Further, the similarity between these patterns and the additional X-slot pattern is completely coincidental. For languages like Guarjío, where the heavy syllable is realized via an epenthetic laryngeal, Raimy could posit the following structure:

(48) pamila ‘boss, governor’

In this case, glottal stop is just an arbitrary consonant to be linked to, “after the first vowel”; there is no way to capture the similarities between the reduplicant shape in this language to the other languages of Uto-Aztecan. In contrast, the constraint-ranking approach, which captures reduplication-specific as well as general historical tendencies in each language, predicts that this kind of variant should be available to some languages.

4.5.3. On the morphological status of the reduplicant

At the outset of this chapter I began with a brief review of different views that have been offered with respect to the morphological status of reduplication. Should it be viewed as Item-and-Arrangement, Item-and-Process, or even an as artifact of a Word-and-Paradigm type grammar?

It is my position that the cross-linguistic evidence of reduplication in Uto-Aztecan suggests that reduplication should be viewed as morphemic pieces, as in an Item-and-Arrangement type morphology. If this is the case, then reduplicative patterns are available to investigation by the comparative method, and are also amenable to historical reconstruction. This position is suggested by Langacker (1977a) with respect to
reduplication in Uto-Aztecan, and I argued above that a heavy syllable reduplicant can in fact be reconstructed for the Proto-Uto-Aztecan distributive morpheme.

Bybee, Perkins, and Pagliuca (1994) also discuss the diachronic development of reduplication, from the perspective of grammaticalization theory. Their focus is on the development of grammatical (i.e. functional) morphemes from lexical morphemes through historical paths brought on by semantic and phonological erosion. Reduplicative morphemes challenge “the principle that all grams develop from a fuller lexical source, since it is not possible to trace a reduplicative gram back to a single word or even a specific phrase” (p. 166). However, Bybee et al. suggest that their theory can in fact be applied to reduplicative morphemes, if we consider the “fullest, most explicit form of reduplication, total reduplication, to be the originating point for all reduplications, with the various types of partial reduplication as reductions and thus later developments from this fullest form” (p. 166). Haugen (2003) critiques this particular aspect of the Bybee et al. approach, pointing out that since reduplicative morphemes in Yaqui do not have a one-to-one form-meaning correspondence, they therefore show that it is not necessarily the case that the grammatical development of reduplicants is evident at every stage of a language. As should be evident from the discussion in section 4.2, this lack of form-meaning correspondence among reduplicants is not uncommon in reduplication in Uto-Aztecan more generally.

However, under the assumption that reduplicative morphemes are pieces of prosody in the underlying representation, it is hoped that future research will be able to add more comprehensive comparative data from the Uto-Aztecan languages in order to reconstruct
the reduplication system of Proto-Uto-Aztecan, from which the paths of grammatical change could then possibly be inferred. In addition, the elaboration of the phonological and semantic changes and shared retentions involved in Uto-Aztecan reduplication might also contribute to the ongoing debate regarding sub-branching within the family. We seem to be at a point where Uto-Aztecan evidence can also be assembled, analyzed, and compared with the results of investigation into reduplication and historical linguistics in other language families, e.g. Austronesian (Bybee et al. 1994, Finney 1999, Kennedy 2003).

4.6. Conclusion

In this chapter we have considered the major theoretical approach to reduplication within the piece-based theory of Distributed Morphology: the modular-derivational theory of Raimy (2000). I have argued that three crucial aspects of Raimy’s theory are undesirable in considering the comparative reduplication patterns evident in Uto-Aztecan, which illustrate cognate patterns of prefixal reduplication, of which we have focused on marked heavy syllable reduplication and mora affixation. From the complications necessary to produce a modular-derivational account of the comparative Uto-Aztecan reduplication patterns, it seems clear that the OT-style approach best accounts for the comparative data. As far as I can tell, excepting the issue of prosodic structure, which I take the Uto-Aztecan data to amply justify, the issue of which theory one adopts is largely a matter of theoretical ideology, and it is not clear what kind of empirical data would lead to a definitive choice between the two. However, OT seems to provide a more constrained
and independently-motivated theory of the variation we see in the Uto-Aztecan reduplication patterns.

I have also argued that reduplicative morphemes should be considered to be prosodic pieces, and in the next chapter we will consider the interaction of reduplicative morphology with syntax, where I will discuss a way to integrate the morphophonological aspects of Optimality Theory with the morphosyntactic aspects of Distributed Morphology.
CHAPTER 5
ON REDUPLICATION AND SYNTAX

5.1. Introduction

This chapter addresses the interaction of reduplication with syntax. The central issue that this chapter seeks to address is how the piece-based analysis of reduplication discussed in Chapter 4 is available to Distributed Morphology, at least with respect to the data that we seen in Uto-Aztecan. I will also address the crucial issue of how a constraint-based theory of phonology such as Optimality Theory is compatible with DM, and how the two can be combined for a theory of interaction between morphology and syntax, on the one hand, and morphology and phonology on the other.

This chapter is organized as follows. Section 5.2 discusses two recent papers by Lisa Travis addressing the question of the position of reduplication in syntax. In section 5.3 I discuss my own proposal for the interaction of reduplication and syntax, and section 5.4 concludes.

5.2. Travis on the syntax of reduplication

Despite the relative paucity of theoretical discussions about reduplication and its relation to syntax, Travis has directly addressed this issue in two recent papers. Travis (1999) presents a preliminary report of her research into the interaction of reduplication with syntax. Although the proposals made in the earlier paper are slightly different from the more recent proposals in Travis (2001), Travis (1999) raises several interesting
propositions that I will consider with respect to the Uto-Aztecan data that we have been examining. Since these papers are not yet well-known I discuss them at length.

5.2.1. Travis (1999): “A syntactician’s view of reduplication”

Travis (1999) begins her discussion by pointing out that reduplication is usually the domain of phonologists and morphologists, but as theoretical morphology has shifted to syntactic models, as with DM,¹ and since reduplication is a particularly salient environment in which to study word-internal structure, reduplication should become of greater interest to syntacticians than has been the case heretofore. Following the reduplicant-as-affix approach of Marantz (1982), Travis claims that a reduplicant is an affix inserted in syntax. Following the tradition of OT I will refer to this reduplicative morpheme as RED.

An example of a syntactic derivation with RED, from Tagalog, is given in (1) (this tree makes use of Travis 1992’s distinction of “Inner” and “Outer” Aspect—the reduplicant marks “incomplete aspect”):

---
¹ As an example of syntactic approaches to word structure, Travis points to the Mirror Principle of Baker (1985), which claims that morphological derivations must directly reflect syntactic derivations, and vice versa. Although this principle was originally intended to act as a kind of constraint holding between two different levels of representation, Morphology and Syntax, in a truly syntactic framework of morphology the “Mirror Principle” is simply a consequence of the syntactic nature of morphology.
In this tree, RED surfaces as the Spell-Out of a syntactic head, in this case, Inner (i.e. vP-internal) Aspect. This differs from trees that we will see below, where RED is the head of its own phrase in syntax.

Contra Marantz (1982), who viewed reduplicants as only minimally different from other kinds of affixal morphemes, Travis suggests that reduplicants may in fact differ both formally and functionally from other affixes. One difference in form between RED and other morphemes is the fact that RED gets its phonological content from other morphemes, depending on what constitutes the target of reduplication. For example, in Tagalog verbal reduplication, RED can target either the verbal Root itself, or the second syllable of the complex stem, as in (2) (examples from French 1989; see also Rackowski 1999):
According to Travis, “without even knowing all the details, the fact that there is variation in position [of the reduplicative target] clearly raises the question of whether reduplication can tell us anything about syntax” (p. 318).

As to the issue of the function of reduplication vis-à-vis other affixes, Travis states that “many cases of reduplication can be seen as quantificational and that these cases of reduplication do, in fact, link form to function and that this function of reduplication is represented in its syntactic position which is encoded in UG in some way” (320). The view that reduplication is (often) quantificational has also been presented by Jelinek and Demers (1997).

In addition, Travis points out that that reduplication is not merely iconic since there are many examples of non-iconic uses of reduplication. For example, two “instances” of the first syllable of a nominal root in Hopi does not indicate “two of that thing”, but
rather, plural (e.g. saqa ‘ladder’ → sasaaqa’ ‘ladders, more than 2’). Dual, on the other hand, is indicated by the suffix -vit (e.g. maana ‘girl’ → maanavit ‘2 girls’) (data from Travis 1999: 320).

Travis concludes that “the position of reduplication may be used to indicate the placement of a syntactic head”, only if “reduplication behaves like (most) other morphemes—that is, [if] it appears in a fixed position that can tell us something about syntax” (p. 317). Travis proposes two hypotheses regarding the syntactic placement of reduplication. First, she proposes that quantifying reduplication is added in the syntax. Second, she proposes that where reduplication occurs in the syntax will determine its form and meaning.

Regarding the first proposal, Travis points out that in addition to the phonological domains targeted by reduplication (e.g. the prosodic hierarchy given in McCarthy and Prince 1986), reduplication can also target syntactic domains. She proposes having reduplication added in the syntax, and “from this syntactic position may target either a syntactic or a phonological domain” (324).

On the issue of “targeting a phonological domain”, it should be pointed out that most work in prosodic morphology has assumed that reduplicants do not actually target a phonological domain, as witnessed by the controversy over the issue of “syllable-copy” reduplication first raised in Moravesik (1978) and taken up in later work, such as McCarthy and Prince (1986) and much work thereafter (see discussion of this issue in Chapter 4). Simply put, the hierarchy of prosodic units shows what reduplicants can be, rather than what they may copy, at least under standard assumptions.
However, it is not clear to me that such targeting of prosodic units is impossible, and in fact I have argued that the first syllable is targeted for reduplication in many (though not all) roots in Yaqui (Haugen 2003). This approach makes a distinction between reduplicative target, a morphosyntactic notion, and reduplicative base, a morphophonological notion. In section 5.3 I will discuss this distinction in greater detail.

Returning to the issue of the multiple reduplicative targets in Tagalog, which are indicative of different duplemes in the terms of Spaelti (1999), Travis proposes a possible link between reduplicative form and semantic function, the latter being linked directly with the syntax. This is the basis of the Form/Function Mapping Hypothesis:

(3) **Form/Function Mapping Hypothesis (F/FMH) (Travis 1999: 323)**

The form of a (reduplicative) affix will mirror the function of a (reduplicative) affix.

The F/FMH applies only to those cases where different reduplicative forms take different meanings, i.e. it only applies to duplemes.

I will consider this claim only in its relativized form, where form and function are language-specific, since the strongest version of this hypothesis (which would call for universal form-function mapping) is obviously false—Travis herself gives counterexamples (see Travis 1999: 324).

Travis states:

The claim that I would like to make is that the form that reduplication takes can determine its meaning. This certainly would . . . lighten the acquisition load on the language learner. The idea is that the higher in the syntactic tree the reduplicative morpheme is, the larger the domain that it targets. For example, a reduplicative
morpheme closer to the root will reduplicate a syllable while a reduplicative morpheme further from the root will reduplicate a foot, still further a word and still further a phrase (323).

This claim is contradicted by the evidence from Uto-Aztecan surveyed in the last chapter, which suggests that reduplicative morphemes, even when “duplemic”, can be stem-specific, and there is no correlation of reduplicant size and meaning that would fit this particular conception (for example, see the different shapes for the expression of habitual in Yaqui).²

The F/FMH is also contradicted by variable reduplication in Yaqui, where some roots can take different yet synonymous reduplicants. For example, the habitual of the complex form *tekipanoa* (> Nahuatl teki- ‘work’ + -panoa ‘to pass’, Dedrick and Casad 1999) can either be *tekipapanoa*, with reduplication of the first syllable of the verb root, or *tekipannoa*, with gemination in the first syllable of the verb root (Constantino Martínez, personal communication). Both of these reduplication forms yield a reading of habitual action, thus showing that the same root can have variable reduplication patterns for the same semantic function. (See also Harley and Amarillas 2003).

In addition to being relativized to language, reduplication may also be relativized with respect to syntactic category. For example, Travis shows that formation of intensive verbs in Tagalog involves a syllable (*lakad* → *maglalakad* ‘walk repeatedly’), whereas adjectival intensification involves a foot (*gutom* → *magpakagutomgutom* ‘try to become

² It should be admitted, however, that I have not examined the possibilities of phrase-level reduplication in UA, focusing instead on grammatical reduplication of words. The fact that no discussion of UA reduplication highlights phrasal reduplication may be an indication that it is not as productive in UA as it is, for example, in Tagalog.
very hungry'). Travis proposes a syntactic tree where "the reduplicative affix that intensifies the verb root is internal to the extended projection of the verb" and "the reduplicative affix which intensifies the adjectival root is at the outermost edge of the extended projection of the adjective" (329). The tree in (4) is her proposed structure for deadjectival verbs in Tagalog:

(4)  

```
vP
  SPEC  v'
    v   REDP
      RED  VP
        SPEC  V'
          V  REDP
            RED  AP
              NP  A'
                A
```

(Travis gives evidence for the internal status of the NP subject, having to do with agreement and case, which is beyond the scope of our interest here). Presumably it is the shape of the reduplicants which Travis takes as evidence of their syntactic position, given the underlying assumptions behind the proposal of the F/FMH.

However, from the Yaqui data in Chapter 4 it can be seen that aspectual reduplicants of that language, which are presumably of the "outer" variety, can surface in a variety of forms, from copy of a full CVCV Root, to a CV syllable, to a bare mora affix. I take this to be evidence that Travis' F/FMH is falsified, and I will not consider it further.
What we can take away from (4) is Travis' proposal to derive a syntactic source for reduplicative morphemes by means of a specific projection, "REDP", in syntax. What I will propose in section 5.3 below is something very different: that RED is simply the Spell-Out of specific features manipulated by syntax, rather than requiring its own phrasal projection. Before turning to that discussion, however, we will examine Travis' more recent proposals regarding the syntactic basis of ostensibly different reduplication types.

5.2.2. Travis (2001): "The syntax of reduplication"

Travis (2001) makes a distinction between three different types of reduplication, with an allusion to a possible fourth type that would warrant further investigation. Travis intends to have each of the types be determined by their underlying syntactic structure. Only two of these are relevant to us here, however: **phonological reduplication** and **syntactic reduplication**.³

Phonological reduplication is the most familiar and most-often studied type of reduplication, and is so-called because it is "sensitive to phonological domains" (p. 456). Travis proposes that phonological reduplication "represents the syntactic structure that is created by head movement and the fact that it is phonological can be predicted by the syntactic structure in which it appears" (p. 256). Following the reduplicant-as-affix proposal of Marantz (1982), and assuming a version of DM, Travis claims that the fact

³ The third type of reduplication is **contrastive reduplication**, of the kind represented by such English examples as *I don't LIKE HIM like him* (Ghomeshi, Jackendoff, Rosen, and Russell 2000). The fourth type is **conjunction reduplication**, as discussed for echo reduplication in Kannada by Lidz (1999). Since neither of these types of reduplication seem to have analogues in Uto-Aztecan, we will ignore them here. The point of mentioning them is that a Travis-style analysis has to posit different syntactic mechanisms to distinguish these types of reduplication from the two kinds that we will be discussing here.
that reduplication appears as an affix is predicted by the model of DM, since this kind of head movement is the typical way to derive affixation in DM.

An abstracted tree structure for phonological reduplication, showing head movement of X to Q, is given in (5):

(5) **Phonological Reduplication**

![Tree structure for phonological reduplication](image)

Travis gives the following tree structure for the plural reduplication form for *ban* ‘coyote’ in Tohono O’odham.\(^4\)

(6) \(ban\) ‘coyote’ \(\rightarrow\) *baa-ban* ‘coyotes’  \(\text{Travis 2001: 457 [3]}\)

\[^4\] Travis actually gives the form *bana* ‘coyote’, taken from Moravcsik (1978), in turn taken from Langacker (1972). It is not clear if this form is a dialectal variant or a typo in the original source, but *ban* is the form cited by Saxton, Saxton, and Enos (1983), and this is the form that I include here.
Travis assumes that "phonological reduplication occurs when the sister of the reduplicative head adjoins to it", and then "a subpart of the sister of the reduplicative morpheme will be copied and will appear in the affix position itself" (p. 457).

It appears, however, that the head movement cannot be the entire story to the account of phonological reduplication, since head-movement is not the only way to induce sisterhood among two morphological items. That is, a sisterhood relation, at least in morphological terms, can also be brought about via Merge. Nevertheless, Travis' notion of a morphological sisterhood relation between the reduplicant and its target seems to be correct, regardless of the syntactic mechanisms necessary to induce this relation.

However, Stonham (1994) gives compelling evidence from Nitinaht (Wakashan) that reduplication can be triggered on a root by an affix that appears non-adjacent to that root. We will return to this issue in section 5.3.

Travis' second type of reduplication is *syntactic reduplication*. Travis claims that syntactic reduplication is similar to phonological reduplication, but she also endeavors to distinguish the two based on the differences in underlying syntactic structure. Following Pi (1995), Travis points out that, like phonological reduplication, syntactic reduplication forms can target different domains, as in (7):

(7) a. cup after cup of coffee
    b. cup of coffee after cup of coffee
    c. cup after steaming cup of coffee
    d. steaming cup after steaming cup of coffee
    e. steaming cup of coffee after steaming cup of coffee
In addition to similarity in form, Travis follows Moravcsik (1978) in positing that syntactic reduplication parallels phonological reduplication in function—i.e. the expression of some quantitative [Q] feature.

The abstract notion of syntactic reduplication is given in the following representation:

(8) Syntactic ("Spec-filling") reduplication

```
  QP
 Spec  Q'
COPY Q  XP
  X  ZP
 Y  X
```

In phonological reduplication, head movement occurs before copying, as this head-movement sets up "the appropriate environment for the reduplication" (p. 461). In syntactic reduplication, "feature checking does not force movement to the Spec position. Rather, the head first creates a copy which will check the feature as shown in ([8])" (461). Thus, in syntactic reduplication the feature-checking occurs only after copying.

There are several facts which Travis claims follow from her distinction between these two kinds of reduplication. As phonological "pieces", affixes have inherent shapes. Specifiers, on the other hand, do not, thus, unlike phonological reduplication, there is no prosodic restriction on syntactic reduplication.

Also, in phonological reduplication we see "left/right mapping" (i.e. Marantz's 1982 generalization that prefixal reduplicants tend to copy from the left edge of a stem while suffixal reduplicants tend to copy from the right edge). Syntactic reduplication, on the
other hand, "seems to be bottom-up, i.e. it must contain the head of the complement independent of whether it is left-most or right-most in the structure" (p. 462).

Finally, syntactic reduplication is fully recursive whereas phonological reduplication typically only copies once. In Nitinaht, even when multiple reduplication-triggering suffixes are attached to a root only one reduplicant appears on the surface. Stonham (1994) analyzes this as morphological haplology, where adjacent (reduplicative) morphemes may "unify" their features, and different reduplicative templates can be satisfied simultaneously (pp. 53-55). Such evidence suggests to me that syntactic reduplication should be regarded as mere syntactic recursion, whereas phonological reduplication is in fact the realization of an actual morphological exponent (i.e. an affix of some sort, e.g. RED).

To carry this point further, it doesn’t seem at all obvious to me that we need to make what we could regard as "rhetorical repetition" a special kind of reduplication, a term that would be better off left to its more grammatically-oriented sense. Thus, for example, in Nahuatl difrasismos, those metaphorical couplets where two terms are placed adjacent to one another in order to invoke a third concept (e.g. in atl, in tepetl, literally meaning ‘the water, the mountain’ but referring to ‘the city’), we need not consider the repeated, phrasal material to have any special status with respect to syntax. Bright (1990) points out that “such metaphorical phrases can occur in a variety of morphological contexts, both inflectional and derivational” (p. 440), such as (9):

(9) in a-hua-que, in tepe-hua-que (Bright 1990: 440 [4])
the water-owner-PL, the mountain-owner-PL
‘the lords of the water, the lords of the mountain’ = ‘the lords of the city’
Bright concludes that “this suggests that difrasismos, although they may represent semantically specialized constructions, need not be given special grammatical status” (p. 440), i.e. need not require special mechanisms (e.g. a special kind of syntactic reduplication) in order to explain the link between their form and meaning. This position is essentially the same as that which is entailed in Marantz (1997)’s view of idioms: the interpretation of lexical items is contextualized to particular syntactic contexts. While in a-tl means ‘the water’ (morphologically it is DET + water-ABS, the Uto-Aztecans “ABS” being the marker of a non-possessed noun), in the context of appearing before in tepe-tl it means ‘the city’, just as cat in the context of the English idiom Let the cat out of the bag means ‘a secret’.

Similarly, grammatical elements can be repeated across a discourse for rhetorical effect in English. For example, in A Tale of Two Cities, after Doctor Manette’s memoir, written during his captivity in the Bastille, is read at Evrémonde’s second tribunal, thus condemning him, Dickens describes the jubilation of the audience (and the Tribunal itself) thus: “At every juryman’s vote, there was a roar. Another and another. Roar and roar.”

Should we consider these repetitions to be simply that, an option employed, here dramatically, by the author given the possibilities made available by a generative system of syntax, or should we consider these some form of grammatical reduplication? Given the lack of systematicity of such examples, I prefer the former analysis.

To conclude this section, if Travis’ conception of syntactic reduplication is indeed different from phonological reduplication, then the types of reduplication that we have
been examining in Uto-Aztecan are of the latter type. It remains to be demonstrated that this is limited to head-movement contexts, however.

Following the Pervasive Syntax Perspective outlined in Chapter 3, I assume that morphology is built up by syntax and then handed over to phonology at Spell-Out, where Vocabulary Items are inserted into appropriate syntactic slots. While Travis’ distinction draws out the importance of reduplication in phrasal syntax, I will focus in the next section on the syntax of word-level reduplication, that what Travis regards as *phonological reduplication*, the traditional domain of phonology-morphology interaction.

5.3. On the syntax of “phonological reduplication”

One crucial aspect of the previous discussion of Travis’ work on the syntax of reduplication is the notion of reduplicative target. Here I define target as the morphosyntactic domain to which reduplication applies. I follow Travis, and much other work, in assuming that this target is in a morphological sisterhood relation to the reduplicant (RED), which I assume to be an affixal “piece” inserted into Morphological Structure at Spell-Out. This relation can be schematically diagrammed as follows, where the target of reduplication is a hypothetical stem PATTA:

\[
(10) \quad X
\]

\[
\begin{array}{c}
\text{[RED]} \\
\text{[PATTA]}
\end{array}
\]

We will discuss the insertion of such a stem into a syntactic derivation below, but first it will be necessary to distinguish this syntactic notion of target from the phonological
notion of base—that morphophonological constituent from which the reduplicant copies to fill in its own phonological material.

Here I assume a version of Correspondence Theory, an adaptation of the “Full Model” of McCarthy and Prince (1995). The original Full Model, which posits three relations crucial to understanding phonological representations, is given in (11):


\[
\begin{array}{c}
\text{Input} \\
\text{Reduplicant} \\
\text{Base}
\end{array}
\]

In this model, there are three possible phonological relations, only two of which will be under consideration here.\(^5\) The first relation is construed by McCarthy and Prince (1995) as the Input-Base, or IB, relation. This is the standard relation of an underlying phonological representation to its surface form. In the constraint-based framework of Optimality Theory, the output form is the candidate that surfaces which best satisfies the series of ranked constraints in a given grammar. Among the constraints are those that consider well-formedness (i.e. Markedness constraints), and those that consider faithfulness to the underlying form (FAITH, e.g. “DEP” and “MAX”).

The Base-Reduplicant, or BR, relation has to do with the correspondence between the segments of the reduplicant and those of the Base. In this model reduplicants do not actually perform an operation of “copying”, but, rather, constraints identify the extent to

---

\(^5\) The third relation, the Input-Reduplicant relation, is controversial, but Fitzgerald (2000) has used evidence from reduplication with syncope in Tohono O’odham to argue that it exists. Even if so, cases where IR relations need to be specified are rare and we will not consider further examples here. (Fitzgerald claims that it is required for Tohono O’odham reduplication because of the peculiar property of syncope in the base of certain forms—see 4.4 for further discussion).
which the segments in the reduplicants of various possible outputs stand in correspondence with (i.e. are "faithful to") the segments of their Base.

An example of how these relations work is given in (12), where we will assume for the sake of argument that reduplication is total:

(12) \[ /\text{PATTA}/ \]
\[
\text{PATTA}_R \quad \text{PATTA}_B
\]

However, the nature of what the "Base" actually is remains under-theorized, as work within OT generally assumes that the entire segmental and featural content of a given stem serves as the "Base" for reduplication. This assumption is encoded in the conflation of the Input-Base (IB) relation with the input-output relation above in (11). Since I do not presuppose that the Base is the complete set of phonological material in correspondence to the Input, as assumed in the McCarthy and Prince (1995) version of the Full Model, Haugen (2004) proposes the following modification:

(13) \textbf{The Modified Full Model}

\[
\text{INPUT}
\]
\[
\text{Reduplicant} \quad \text{[OUT]}_R\text{-PUT} \quad = \quad \text{Surface Form}
\]

In this version of the Model, the segments corresponding to the underlying Input segments, e.g. PATTA, are not necessarily all part of the reduplicative Base, which I have enclosed in brackets with a subscripted B indicating that \textit{only} these elements comprise the base. There must be a BR relation, since a Base needs to be defined in order to provide the phonological material with which the segments of the reduplicant must be in correspondence, but in my view this Base need not correspond entirely to an
input stem. The "Surface Form" is the entire structure composed of the Reduplicant, the Base, and any material left over in the Output—in short, the Surface form is composed of the Reduplicant plus the Output. A concrete example is given in (14), where the Base is only the first syllable of the Output form:

(14) PATTA

\[
\begin{array}{c}
\text{PAT} \\
\downarrow
\end{array}
\quad [\text{PAT}]_{n-TA} \quad \text{yields} \quad [\text{PAT}].[\text{PAT}].TA \rightarrow [\text{pat}.pat.ta]
\]

In (14) the reduplicant is once again a full copy of the Base, although it is not a full copy of the Output (i.e. the entire stem corresponding to the underlying input).

My approach does raise the crucial issue, left tacit in most discussions of reduplication in the (OT) literature, of how the Base gets decided: if the Base is not the full input stem, what is it? Perhaps a more crucial question is why I have taken this approach and not just assumed, along with most authors, that the Base is in fact all of the stem. To answer the second question first, recall that Uto-Aztecan marked heavy syllable reduplication typically does not copy into the second syllable of the stem, as shown in (15):

(15) a. Yaqui: yena \rightarrow yey.yena ‘smoke from time to time’
    b. Mayo (accented): nóka \rightarrow nón.noka ‘know a language (hab.)’
    c. Guarijío: pamila \rightarrow pa?.pamila ‘boss (pl.)’
    d. Tohono O’odham: ?ag \rightarrow ?a?.?ag ‘animal horns (distr.)’
    e. Tohono O’odham: ban \rightarrow baa.ban ‘coyote (pl.)’
    f. Nahuatl: o:me \rightarrow oh-o:me ‘two at a time’
    f. Comanche: ten.séé \rightarrow téh.ten.séé ‘ten cents/ ten cents apiece’

Under the Base=Stem approach, there is no reason not to expect copy into the second syllable, as with the unaccented class in Mayo (e.g. noká \rightarrow nok.nóka ‘speak

\text{\footnotesize\textsuperscript{6}} As above, I remain agnostic about a possible IR relation, which is not relevant here if it does exist. The reconciliation of this model with such a relation remains a worthy topic for future investigation.
(habitually)' and as is expected in other languages (e.g. most Austronesian languages). If the base is only the first syllable, then it is fully copied in cases of CV.CV- stems with heavy syllable reduplicants, but the heavy syllable requirement must be created through some other form of mora augmentation: usually gemination, but also laryngeal-epenthesis or vowel-lengthening. Defining the base in some way is also necessary to capture languages, like Mayo, that have different bases for different classes of words (see Hagberg 1993 and section 4.2.1.2 of this dissertation for argumentation supporting this point).

As to the issue of how the base gets limited to the first syllable, there seem to be two possible approaches. The first is unattractive from the standpoint of evaluating the theory, since it involves simple lexical-specification: stems of X class have base of a shape (e.g. a syllable), and so on. Under this approach, reduplicative bases would be specified in the lexical entry for each root. This certainly gets the job done, but such an approach violates notions of economy and the underspecification of underlying representations.

A second approach would be to define the reduplicative base by means of constraints, such as \( B_C = \sigma \) (read: "bases in class C = a syllable"). Haugen (2004) treats the Mayo reduplication patterns in this way, and this approach to the Mayo data requires a suite of alignment constraints to capture the variant stress patterns, base sizes, and reduplicant

---

7 I say most here because the paradigm cases of heavy syllable reduplication discussed by McCarthy and Prince (1986, 1993) involve such Austronesian languages as Ilokano. However, Ballantyne (1999) has independently argued for Yaqui-style syllable-copy reduplication in the Micronesian language Yapese. Crucially, in both Yaqui and Yapese coda consonants are non-moraic, so copying them in reduplication does not lead to a violation of reduplicant well-formedness conditions (e.g. the equivalent of light syllable-
sizes in the two major classes of Mayo words. Along with Shaw (2004, in press), I assume that bases defined by constraints would be limited to either prosodic units (e.g. syllables or feet) or morphological units (e.g. roots or affixes); this position is identified as the "Constituent Base Hypothesis" (CBH) by Shaw (2004, in press).\(^8\) In addition, the approach of specifying bases for certain classes of words should leave open the option of default base-assignment to the entire stem.

Returning to the issue of reduplicative targets, as has been pointed out by Jelinek and Demers (1997), Travis (1999), and others, reduplication is often (though not necessarily always) the expression of some quantificational feature. However, unlike Travis (1999), I do not propose that reduplication heads its own phrase. More in line with Travis (2001), which frames reduplication as the expression of a quantificational (Q) head, I propose that reduplication is simply inserted as an abstract Vocabulary Item, RED, into a particular syntactic structure in order to fill in the phonological (p-) features of some morphosemantic feature(s). From the perspective of DM, which posits the principle of Late Insertion (see Chapter 3), these morphosemantic features are the primitive units that are manipulated in syntax. It is only at the level of Morphological Structure that phonological material, including the abstract reduplicative/prosodic Vocabulary Item RED, is inserted to spell out these features.

For example, in order to spell out the aspectual notion of habituality in Yaqui reduplication, RED would be inserted to spell out the head of Asp(ect)P, as is shown,

\(^8\) The possible bases given by Shaw (2004) are MCat (Morphological Word, Stem, Root), PCat (Prosodic Word, Foot, Syllable, Nucleus, Mora), PHead (L/R-most Foot=PrWdHead, Nuc=oHead, etc.), and CanonCat (Optimal Root = [CVC], Optimal Stem = [CVCV]).
with some abstraction of VP-internal syntax, in the derivation of a typical unergative verb in (16):

(16)a. \[ \text{AspP} \]
   \[ \text{vP} \]
   \[ \text{Asp} \]
   \[ \sqrt{\alpha} \]
   \[ \gamma \]
   \[ \beta \]

b. \[ \text{AspP} \]
   \[ \text{vP} \]
   \[ \text{Asp} \]
   \[ \sqrt{\gamma} \]
   \[ \beta \]

c. \[ \text{AspP} \]
   \[ \text{vP} \]
   \[ \text{Asp} \]
   \[ \sqrt{\gamma} \]
   \[ \beta \]

In the transition from (16)a to (16)b we see the incorporation of the Root complement, \(\alpha\), of \(\gamma\), \(\beta\), into \(\gamma\), which derives an unergative verb (Hale and Keyser 1993). (16c) illustrates that this complex \([\beta, [\alpha, \beta]\]) then moves higher in the tree to check its aspectual features in the head of AspP.

At the level of Spell-Out, at Morphological Structure, Vocabulary Items would be inserted as such: \([[[\alpha bwiika] \beta \emptyset] \gamma \text{RED}]\). The hierarchical ordering of the morphemes can be illustrated with a traditional morphosyntactic tree, although it should be borne in mind that the phonological content is not actually spelled-out until after this syntactic structure is sent to MS.

(17)a. \[ \text{AspP} \]
   \[ \text{vP} \]
   \[ \text{Asp} \]
   \[ \sqrt{\beta} \]
   \[ \gamma \]
   \[ \text{RED} \]

b. \[ \text{Asp}^0 \]
   \[ \text{vP} \]
   \[ \text{Asp} \]
   \[ \sqrt{\beta} \]
   \[ \gamma \]
   \[ \text{RED} \]

From MS, which can be construed as the underlying representation of an Optimality Theoretic tableau, the language-specific constraint-ranking will determine the optimal
output in the spelling out of the phonological content of RED: i.e. the interaction of the language-particular constraints will decide which candidate best satisfies MAX-BR.

Given that syntax builds the structure presented to morphology, one might expect that the order of morphemes would be consistent for any particular configuration. Thus, for example, it might seem odd to have habitual aspect in Yaqui realized as prefixal reduplication, while all other aspects are realized by suffixation of particular morphemes: e.g. -taiite 'inceptive', -(e)k 'perfective', etc.:

(18) bwi-bwika 'sing habitually' bwiika-k 'sang (perfective)'

\[
\begin{align*}
\text{a.} & \quad \text{AspP} \\
& \quad \text{vP Asp [+habitual]} \\
& \quad \text{RED-} \\
& \quad \text{v} \\
& \quad \text{bwiika} & \quad \text{v} \\
& \quad \text{Ø} \\
\text{b.} & \quad \text{AspP} \\
& \quad \text{vP Asp [+perfective]} \\
& \quad \text{-k} \\
& \quad \text{v} \\
& \quad \text{bwiika} & \quad \text{v} \\
& \quad \text{Ø}
\end{align*}
\]

Unlike typical morphemes, which get their phonological content at Insertion, RED gets its phonological content only from other morphological material to which it is adjacent, and only in the PF component. However, each reduplicant must carry some information as to its prosodic structure, as with the specification of a prosodic template, defined by McCarthy and Prince (1993) as “Morphological Category = Prosodic Category”. The templatic view of reduplicants has recently fallen under fire in theoretical phonology because in certain languages not all reduplicants are instantiated by actual prosodic units (cf. the discussion of bare consonant reduplication in Hendricks 1999, such as Semai “expressive minor reduplication”, where the form c?e:t ‘sweet’

---

9 The features listed here in Asp are given only as heuristics—they probably involve more abstract notions bundled together to yield what we interpret as “habituality” or “perfectivity”, etc.
reduplicates as $ct-c?e:t$). However, equivalent notations, e.g. the alignment of the reduplicant to the edge of a prosodic unit, could be invoked to derive the correct results.

Under the view adopted here, the location of the reduplicant with respect to its base (or target) is determined by the morphophonology. In some languages, e.g. Nitinaht, reduplication is the only morpheme that is prefixal, all other morphemes being suffixal. Stonham (1994) takes this as evidence that reduplicants are therefore not affixes per se, but rather are the result of a re-adjustment operation triggered by certain suffixes. As Stonham puts it,

Contrary to claims that reduplication is just another form of affixation, in Nitinaht reduplication is, in many cases, tied to both a suffix and to vowel length and as such must be viewed as involving a constraint on the shape of the root required by the suffix and not as simply an affix on its own (p. 60).

As Stonham notes, the constraint on the output form triggered by these suffixes is reminiscent of the templatic morphology triggered by affixes in Yawelmani (Archangeli 1983). Consider the data in (19), where we see the a long vowel reduplicant on the root $\acute{\lambda}uq^w$- ‘wide’ in (19)c, which is triggered by the addition of the suffix $-a:p$ ‘very’:

(19)a. $\lambda'uq^w$-
   ‘wide’

b. ći:+a:?dĄ
   pull-a.long.object
   ‘it was pulled along a long object’

c. $\acute{\lambda}u:\lambda'uq^w+a:?$dĄ+a:p
   RED-wide-along.a.long.object-very
   ‘X’s legs are too big’
The long vowel reduplicant in (19)c is required by the addition of the suffix \(-a:p\), even with an intervening suffix \(-a: di\), which has no such effect (cf. 19b).

Data from languages like Nitinaht present a puzzle for my account of reduplication, since suffixal and prefixal elements generate double exponents for a given morpheme. Under my analysis the reduplicant itself must be regarded as a morphemic piece, and so in these cases we do see double exponence. Of course, double exponence is not unattested in the world’s languages, occurring, for example, with reduplication and suffixal morphology for pluralization in Tümpisa Shoshone, e.g. \(pe.tü\) ‘daughter’ yielding \(pep.pe.tü-m.mü\) ‘daughters’ (Dayley 1989). For the cases of Nitinaht, I would be interested to know if some explanation for the double exponents with reduplication and suffixes is to be found in the diachronic development of these constructions.

Data like those in (19) above show us that a reduplicant and its base need not be (structurally) adjacent. What I would like to suggest here, perhaps tentatively, is that a reduplicant and its target must be. A target can be a morphologically complex entity, of which the base may be only a sub-constituent.

Returning to discussion of reduplication in Uto-Aztecan, in (18) the target for reduplication is the complex head \([\sqrt{v} + v]\), which is semantically necessary to capture the habitual interpretation. Other possible targets that could be involved would be the root alone \([\sqrt{v}]\), and also \(v\) alone. In addition, some suffixes may also reduplicate, e.g. the Yaqui example \(bwik-tai-taite\) ‘always starting to sing’, indicating the spelling out of Q features in the relevant verbal projection, rather than on the verbal root itself.

---

10 Shaw (in press) gives other examples of reduplication where the reduplicant is not adjacent to its base.
The prefixal or suffixal nature of the reduplicant itself, however, is largely a matter left to the phonological component, wherein alignment and anchor constraints determine the location of the reduplicant with respect to its target, as is generally assumed in the OT literature.

To summarize, although the reduplicant and its target are morphosyntactic entities, the actual phonological realization of the reduplicant is a question left to phonology, as is whether or not the reduplicant is a prefix or a suffix. Thus, the account sketched here makes no predictions about the semantics of reduplication with respect to reduplicant size (light or heavy syllable, etc.), or whether or not the reduplicant is a prefix or a suffix. The only thing that matters for syntax is where the reduplicant is inserted, and what its target is. Similar to what is proposed in Travis (2001), the derivations of the unergative verbs discussed above involve the creation of a sisterhood relation between a reduplicant and its target brought about via head-movement. However, I also assume that this sisterhood relation could also be brought about via Merge (i.e. base-generation). The target subsequently serves as the domain in which the base is assigned in the PF component, possibly by means of alignment constraints. In accordance with the tenets of Optimality Theory, the optimal phonological expression of the reduplicant is selected among competing candidates by the interaction of the markedness and faithfulness constraints in a given grammar.
5.4. Conclusion

In this chapter I have reviewed two recent proposals from Travis (1999, 2001) relating reduplication to syntax. I have followed Marantz (1982) and subsequent researchers in claiming that reduplicative morphology is affixal. I have also claimed that the base for reduplication, a morphophonological entity, can be a subset of the morphosyntactic entity that I have called the target, and that the reduplicant that surfaces for a given morpheme in a given language is determined by the language-specific ranking of markedness and faithfulness constraints. Once a RED morpheme is inserted at Morphological Structure, its phonological realization is entirely a matter left to phonology (PF). This concludes our discussion of prosodic morphology in Uto-Aztecan.
CHAPTER 6
NOUN INCORPORATION AND DENOMINAL VERBS

The term 'incorporation' has been much used in discussion devoted to the structure of American languages. Despite the steadily growing mass of American linguistic material, a good share of the data presented in the last few decades being distinctly superior from the point of view of critical analysis to much that served as illustrative material in earlier days, it can not be asserted that the term is always clearly understood or satisfactorily defined. (Edward Sapir, 1911: 250)

6.1. Introduction

As the epigraph to this chapter suggests, the correct application of the term 'incorporation' had been in dispute up to the time of Edward Sapir's writing in 1911, and I will show in this chapter that the term is still used in different ways by different scholars using different assumptions about language and how it works, as well as about how to describe and analyze it. In addition to laying out the theoretical background requisite for our discussion of the morphosyntax of the Uto-Aztecan languages, I would also like to provide here a refining (and re-defining) of the terms to be employed.

For Sapir, noun incorporation was the “process of compounding a noun stem with a verb... no matter what the syntactic function of the noun logically is” (1911: 257). In Sapir’s view, morphology and syntax were to be construed as obviously separate phenomena with their own rules of application. As Sapir elaborates with respect to the
difference between the English sentence *I write songs* and the noun-verb compound in *I song-write*, in the latter case “the grammatical expression of a logical relation, in other words a syntactic process, is sacrificed to a compositional process in which the logical relation is only implied” (257). With respect to the indigenous languages of the Americas, Sapir goes on to suggest that “the sacrifice of syntax to morphology or word-building is indeed a general tendency in more than one American language” (257).

According to Sapir, Kroeber (1909), on the other hand, had defined noun incorporation in such a way as to invoke syntax as well as morphology: “Dr. Kroeber defines [noun incorporation] as follows: — ‘Noun incorporation is the combination into one word of the noun and the verb functioning as the predicate of a sentence’” (254). Sapir criticizes such a definition with the following remarks: “Noun incorporation is primarily either a morphologic or syntactic process; the attempt to put it under two rubrics at the same time necessarily leads to a certain artificiality of treatment” (255).

Although the pioneers of Native American linguistics at the turn of the last century clearly did not have the theoretical tools to link morphology to syntax as overtly as has been done in recent work in generative grammar, and in particular in Distributed Morphology, Kroeber’s (1911) response to Sapir (1911) viewed incorporation as a “process” no different from compounding in Indo-European. Unlike Sapir, however, Kroeber (1911), in his later paper, emphasized the sameness of the process of noun incorporation in Native American languages and compounding in Indo-European languages. For a modern reading of the debate between Sapir and Kroeber, and for
justification for the claim that Sapir was “the first Lexicalist” (and by extension, Kroeber, at least in his 1909 paper, the first “anti-Lexicalist”), see Sadock (1990: 78-82).

The change in Kroeber's own position between his publication of 1909 and that of 1911 is telling. In 1909, Kroeber had actually argued that noun incorporation as a syntactic process could not exist. He states, “Until subjective noun incorporation is established, or a reason shown why it should not occur, there must be an ideal doubt as to the existence of object incorporation” (1909: 573-4). By 1911, in the words of Sadock, Sapir had “won his theoretical point [i.e. that noun incorporation exists and is a process akin to compounding], and Kroeber quickly and completely capitulated... noun incorporation was to be understood as nothing other than the morphological compounding of the noun with a verb” (1990: 81).

Although Kroeber did indeed capitulate to Sapir's view of noun incorporation as compounding, I highlight here Kroeber's universalistic stance toward grammar, in that he recognized that identical processes operate in, or at least are available to, all languages. Thus, even while acknowledging, in “capitulation to Sapir”, that noun incorporation is compounding, Kroeber also concluded that “it is thoroughly misleading to designate the same process respectively ‘composition’ [i.e. compounding—JH] and ‘incorporation’ according as one has in mind his own or other forms of speech” (1911: 584). In addition, presaging the modern notion of Universal Grammar, which is generally attributed to the “Chomskyan Revolution” that occurred some fifty-odd years after the Sapir-Kroeber debate on noun incorporation, Kroeber laments the current state of linguistic theorizing at the turn of the 20th Century by making the following comment:
Some day philologists will approach their profession not with the assumption that languages must differ in kind, . . . but with the assumption that exactly the same fundamental processes run through them all, and with the realization that it is only by starting from the conception of their central unity of type and method that their interesting and important diversities can be understood. (Kroeber 1911: 584)

Returning to the issue of noun incorporation, in this dissertation I will argue that there are two distinct kinds of noun incorporation that must be recognized: the compounding noun incorporation that has been recognized since Sapir (1911), but also the syntactic incorporation which has more recently been argued to exist by such scholars as Sadock (1980, 1986), Baker (1988), and Hale and Keyser (1993, 2002). The coexistence of the latter of these with the former is a more controversial, theory-specific claim, than the mere existence of one or the other.

Although “noun incorporation” (NI), in its various incarnations, has often been simply equated with compounding, this does not negate the possibility that an additional syntactic process also exists—one wherein some verb-formation occurs in the syntax by means of head movement, and particularly by movement of the head of a complement noun phrase (NP) into the verbal head (v). Thus, like Kroeber, I will view NI in light of the processes used, but whereas Kroeber was only in the position to identify, after his “capitulation” to Sapir, the process of compounding, I will also argue that some verbs are formed via the movement of heads in syntax (Baker 1988). The only syntactic processes necessary to derive the empirical results are the Minimalist Program notions of Merge
and Move (Copy), both of which involve the formation of complex heads, but by means of compounding or head movement, respectively.

The issues raised by Sapir and Kroeber are not limited to the compounding word formations found with "Noun Incorporation", however. Identical issues also arise with respect to the mechanisms of denominal verb formation. Following much recent work in the tradition of generative grammar, and in Distributed Morphology in particular, in this dissertation I have adopted the Pervasive Syntax Perspective (PSP) with regard to the derivation of words, holding that the processes involved in the derivation of all words are syntactic in the narrow sense, at least by initial hypothesis. This position is best elaborated in contrast to what we may regard as the Sapir-like Lexicalist stance, which holds that such processes as verb-derivation are phenomena separable from the principles that operate over syntax proper. Before discussing the specific theoretical proposals that I will make regarding derivational processes in Uto-Aztecan in Chapter 8, it will be helpful to first contrast the two prevalent approaches to incorporation and denominal verb formation in the recent literature on these issues: the Lexicalist (i.e. Sapirean) versus syntactic (i.e. neo-Kroeberian)\(^1\) approaches, embodied very clearly in the theoretical debate between Mithun (1984, 1986) and Sadock (1980, 1986).

\(^1\) Although Kroeber (1909) defined incorporation in syntactic terms, he went on to argue, unsuccessfully, that such a process did not, and in principle could not, exist. One reason behind his claim that such a process could not exist is the asymmetry between subject and object incorporation, a lacuna which still plagues Lexicalist accounts of the process but is explained, if it is true, under the formal, Government account of incorporation proposed by Baker (1988)—see further discussion below. I use the term neo-Kroeberian here to refer to the Universalistic framework that the PSP presupposes: the same operations and principles are available to all languages, and differences must be accounted for by noting variation in the implementation of language-specific parameters.
6.2. Denominal verb formation and Incorporation, or Denominal verb formation as Incorporation?

One of the earliest theoretical works to make an explicit link between denominal verb formation and noun incorporation was Sadock's (1980) discussion of denominal verb formation in Greenlandic. Sadock argues that denominal verbs are formed in syntax via the incorporation of nouns into a verbal head in Greenlandic syntax, utilizing the equivalent of the Chomskyian notion of movement of a syntactic head (here, N) from its phrasal position (here, NP) into another head position (V). Evidence that Sadock used for this claim included especially the fact that the incorporated nominal may be modified, even though the noun is formally attached to the verb stem. Similar data will arise in our discussion of Hopi, Yaqui, and other Uto-Aztecan languages below.

However, subsequent scholars, notably Mithun (1984, 1986, 1999), have attempted to separate the two processes, and, like Sapir, claim that noun incorporation is in fact morphological and not syntactic, although Mithun admits that NI “is perhaps the most nearly syntactic of all morphological processes” (1984: 847, emphasis added).

From a functionalist perspective, Mithun (1984) claims that all languages that have NI also have syntactic paraphrases, and as such “it would certainly be inefficient for languages to preserve exactly equivalent expressions so systematically. The fact that

---

2 Sadock himself cites Rischel (1971, 1972) as an oft-neglected precursor who recognized the import of the Greenlandic data that Sadock discusses for theories of morphology-syntax interaction. Alas, this work by Rischel is still largely neglected in the literature, and I begin my discussion here with the elaboration of Rischel's earlier points as made by Sadock.

3 In the framework of Autolexical Syntax (Sadock 1985, 1991), there is no actual movement. Rather, Morphology may map onto Syntactic Structure in a variety of ways. While I am adopting different theoretical assumptions from those of Sadock, what is crucially clear is that Sadock links the morphological structure to the syntactic structure, rather than having an independent level of lexical syntax building morphological pieces to then place into syntax, which in this respect at least makes the approach of Autolexical Syntax consistent with the goals (if not the assumptions and tenets) of DM.
productive morphological constructions of this type never exist in a language without syntactic analogs indicates that the morphologicization itself must be functional” (p. 848). That is, speakers always incorporate for some reason, although the reasons may vary, from language to language and perhaps even from speaker to speaker. For Mithun, this optionality is one of the crucial ways in which NI differs from derivational processes such as denominal verb formation, where certain affixes require a nominal “host” (e.g. mandatory “noun incorporation” in Inuktitut, where light verbs must attach to a noun in order to create a well-formed verb, Johns 2003). In other words, for Mithun the distinguishing characteristic between noun incorporation and a derivational process such as denominal verb formation is that the former involves the compounding of identifiable roots which may also appear in isolation (i.e. non-incorporated), whereas the latter involve affixes which must attach to roots (see also Mithun 1999: 44-56).

In order to critique Mithun’s position on what noun incorporation is, we must first consider her stance on the different types of noun incorporation (section 6.2.1), and the position of noun incorporation with respect to morphology and syntax (section 6.2.2).

6.2.1. Mithun (1984) on the typology of noun incorporation

Mithun (1984) recognizes four kinds of “noun incorporation” (NI), a theoretical notion which for her differs drastically from that recognized by Baker (1988) and subsequent work from generativist perspectives. According to Mithun these four types of NI form an implicational hierarchy, which she suggests is evidence for a trajectory of grammaticalization.
Type 1 NI involves lexical compounding, which in many cases can involve cross-categorial compounding, e.g. noun (N) + N = N; verb (V) + V = V; but also V + adjective (A) = V; V + N = V; etc. Mithun notes that N-V compounding usually occurs for a reason: “some entity, quality, or activity is recognized sufficiently often enough to be considered name-worthy in its own right” (p. 848), and she points out that “the term ‘incorporation’ is generally used to refer to a particular type of compounding in which a V and N combine to form a new V” (848). In such cases the noun bears a specific relation to the verb: patient, location, or instrument. Mithun notes that this is not necessarily true for other kinds of compounds: e.g. in N + N = N compounds, the relationship of the N’s can be vague (e.g. alligator shoes can denote shoes for an alligator, shoes made of alligator skin, shoes that look like alligators, shoes for an alligator costume, etc.). Type 1 NI can be subdivided into two sub-types: compounding by juxtaposition and morphological compounding.

In *compounding by juxtaposition*, independent words can retain their own stress, and therefore the elements involved are still independent words, but they are physically placed next to each other and function syntactically as a unit, as in Samoan clitic placement, where clitics usually follows the V but which follow the VN compound after

---

4 As an aside, Mithun does not address the issue of why the theta-role of agent is unavailable for such “incorporation”—surely it is just as functional for speakers to commonly rehash the frequent agents of activities as the frequent patients, locations, or instruments of those activities. Mithun does not account functionally for the apparent lack of possibility of (agentive) subject incorporation, an empirical observation dating at least as far back as Kroeber (1909). From a strictly formal perspective (e.g. Kratzer 1996, Harley 1995, Marantz 1997a, and others), where agents are projected as specifiers to a causative light verb, this restriction is naturally (if not functionally) explained. [More recent works have attempted to expand the functional projections associated with “verbalness” and causativity. Pylkkänen (1999, 2000), for example, separates Voice (which projects an external argument and is thus the source of the Agent theta-role) from v (the functional head which “verbalizes” a root). The specific details regarding the correct (VP-external) functional projections are orthogonal to my point here, however].
incorporation (Mithun 1984: 850). The placement of the indefinite direct object before
the V in Guarijio OVS sentences may be an example of this type of “compounding by
juxtaposition” incorporation. However, because the incorporated noun does not form a
single phonological word with the verb, Gerdts (1998) regards composition by
juxtaposition, which she terms *noun stripping*, as a process “very much like”, but not
equivalent to, NI (pp. 93-94).

In *morphological compounding*, compounds are single phonological words
undergoing regular phonological processes. Examples of this type come from the Numic
language Comanche (NUA), where NV compounds denote “name-worthy” activities
(data from Canonge 1958, cited in Mithun 1984):

(1) **Morphological compounding in Comanche** (Mithun 1984: 855)
   a. *narînoo’-rîki-
      saddle-put.away
      ‘to saddle up’
   b. *pîki-tsah-kwe’ya-
      fuzz-by.hand-remove
      ‘to skin an animal’

According to Mithun, “because constituent stems often lose their identity under NI, the
meanings of compounds are often not exactly equivalent to the meanings of their parts. In
fact, compounds are often created to convey specialized meanings” (855), as in the
following Comanche examples:

---

5 In Guarijio, when the subject and object are both definite the typical order is VSO or VOS (Barreras
2000).
(2) **Comanche morphological compounding in idiomatic usage** (Mithun 1984: 855)

a. *wana-roh-peti-
cloth-by.force-throw
'to gamble'

b. *waa-hima
cedar.tree-take
'to celebrate Christmas'

For Mithun, this "noun incorporation" as a conjunction of N and V is a morphological process, and is *not* a result of syntax. V + N = V compounds form an intransitive predicate denoting a unitary concept, the name of an institutionalized activity or state.

Type 2 NI involves "the manipulation of case". Whereas Type 1 NI simply lowers the valence of the V when it derives intransitive predicates from transitive ones. . . Type 2 NI advances an oblique argument into the case position vacated by the [Incorporated] N[oun]. When a transitive V incorporates its direct object, then an instrument, location or possessor may assume the vacated object role. When an intransitive V incorporates its subject, another argument may be advanced to subject status. (Mithun 1984: 856)

Mithun gives examples of Type 1 and Type 2 NI for Yucatec Mayan (data from Bricker 1978). According to Mithun, the effect of Type 1 compounding in this language is "to denote a unitary activity, in which the N modifies the type of activity predicated, but does not refer to a specific entity" (857):
Type 1 NI in Yucatec Mayan (Mithun 1984: 857)

a. \textit{t-in-č’ak-0-ah} \textit{čé’}  
\textit{COMP-I-chop-it-PERF} \textit{tree}  
‘I chopped a tree’

b. \textit{č’ak- čé’-n-ah-en}  
\textit{chop-tree-ANTIPASS-PERF-I (ABS)}  
‘I wood-chopped’ = ‘I chopped wood’

(3)b, with the \textit{ABS}-suffix marker \textit{–en} on the verb, is intransitive. According to Mithun, “NI can have a second effect: the direct object role vacated by the IN can be assumed by an oblique argument of the clause”, leading to a transitive verb (marked by transitive suffix \textit{–t}):

Type 2 NI in Yucatec Mayan (Mithun 1984: 858)

a. \textit{k-in-č’ak-0-k} \textit{čé’} \textit{ičil in-kool}  
\textit{INCOMP-I-chop-it-IMPF} \textit{tree} \textit{in my-cornfield}  
‘I chop the tree in my cornfield’

b. \textit{k-in-č’ak-čé’-t-ik} \textit{in-kool}  
\textit{INCOMP-I-chop-tree-TR-IMPF} \textit{my-cornfield}  
‘I clear my cornfield’

In (4)b, where the nominal root \textit{čé’} ‘tree’ is incorporated, the preposition is lost and \textit{in-kool} ‘my cornfield’ becomes a direct object. With the omission of the preposition \textit{ičil}, the cornfield is “affected” in a way that it is not in (4a), in that it is the recipient of the action of the verb.

Another example of Type 2 NI comes from Tupinambá (Tupi-Guarani):
(5) TYPE 1 NI in Tupinambá (Mithun 1984: 856-7)

a. \[a-’i-’u\]
   I-water-ingest
   ‘I drink water’

b. \[a-ka-’u\]
   I-kawi-ingest
   ‘I drink kawi’

c. \[a-ma’e-’u\]
   I-NONHUMAN-ingest
   ‘I eat non-human objects’

d. \[a-por-ú\]
   I-HUMAN-ingest
   ‘I eat human flesh’

(6) TYPE 2 NI in Tupinambá (Mithun 1984: 856-7)

a. \[s-oBá a-yos-ey\]
   his-face I-it-wash
   ‘I washed his face’

b. \[a-s-oBá-éy\]
   I-him-face-wash
   ‘I face-washed him’

According to Mithun, without noun incorporation, the “direct object is the face, marked by the prefix \(-yos-\), as in [6a]; but with NI, the derived object is the owner of the face, marked in the V by the prefix \(-s-\), as in [6b]” (857).

At least from these data it is not clear that Mithun’s conclusions regarding Tupinambá are correct. It looks like (6)b may just be the incorporation of the free NP in (6)a, and it is not clear why (6)b is not glossed as “I-his-face-wash”. This question would be answered with some paradigm in which the possessive modifier was different from the object modifier.
However, as far as I can tell Mithun’s Type 2 NI is almost equivalent to the availability of using pronominal arguments on verbs (Jelinek 1984), or “optional polysynthesis” (Baker 2001). Examples (3), (4) and (6) all show that the verb requires some direct object argument, which can be an inflected noun root or a pronominal element (null for the third person in [3a]), and which can be elaborated upon with an adjunct NP. As Mithun states: “Instead of simply reducing the valence of the V by one, it permits another argument of the clause to occupy the case role vacated by the [incorporated noun]. The result is a lexical device for manipulating case relations within clauses” (p.859).

Mithun’s Type 3 NI involves the manipulation of discourse structure. This NI is used to background known or incidental information, and typically occurs in polysynthetic languages, where V’s contain obligatory pronominal affixes referring to subjects and objects (or agents and patients), and full NPs are optional.

A typical example comes from the SUA language Nahuatl (data from Merlan 1976, cited by Mithun 1984):

(7) TYPE 3 NI in Huahita Nahuatl (Mithun 1984: 860-1)

a. *askeman ti-*'kwa *nakatl
   never you-it-eat meat
   ‘You never eat meat’

b. *na’ ipanima ni-naka-kwa
   I always I-meat-eat
   ‘I eat it (meat) all the time’

---

6 The gloss for (7)b seems misleading. With the root for meat incorporated into the verb, it seems like it should be translated as ‘I eat meat all the time’, contra (what I presume to be grammatical: ni-*kwa ‘I-it-eat’, parallel to 7a). This is in fact how this sentence is glossed in the original source (Merlan 1976: 185).
In the parlance of generative grammar, Mithun’s Type 3 NI seems to be equivalent to Obligatory Pronominal Arguments (or Obligatory Polysynthesis) (Baker 2001, but see Chapter 9)—the verb must be inflected with reflexes of agent and patient arguments, reflected either in the form of pronominal arguments or by incorporated nominal roots. While verbs must always have (minimally) their pronominal arguments, full NPs are also available for speakers to manipulate in discourse contexts (see also Jelinek 1984).

Finally, Mithun’s Type 4 NI involves what she terms classificatory noun incorporation, where a generic N is incorporated to qualify the V, while a more specific external NP identifies the implied referent. An example is from Caddo (Caddoan), where the root –‘ičah- ‘eye’ is used as a classifier of small round objects, “instead of just kinds of eyes” (865):

(8) Type 4 (“Classificatory”) NI in Caddo (Mithun 1984: 865)

a.  
  kassi’ hâh-‘ičá-sswi’-sa’
  bead PROG-eye-string-PROG
  ‘She is stringing beads’

b.  
  ka’as hâh-‘iča-‘i’-sa’
  plum PROG-eye-grow-PROG
  ‘Plums are growing’

For Mithun, these NI types are not random, but instead form an implicational hierarchy: Type 1 >> Type 2 >> Type 3 >> Type 4. Once a language is able to do Type 1, it is on a path to potentially do the subsequent types as well, and a language can only get to Type 3 after going through Type 2, etc. According to this hierarchy, for a

---

7 This is not a necessary trajectory: “decay” in productivity can occur at any point. Mithun gives the following diagnostics for decay in each of the types:

Type 1: diminishing numbers of increasingly opaque compounds;
language to have Type 4 NI, that language must also have Type 3 NI. That is, classificatory NI should only occur in polysynthetic languages, in the terms of Baker (1996). In Chapter 7 we will see empirical evidence from across Uto-Aztecan that this implicational hierarchy is incorrect: classificatory NI can occur in non-polysynthetic languages (i.e. in languages that do not have obligatory object-marking on their verbs). In Chapter 9 I claim that polysynthesis actually developed in Nahuatl from obligatory classificatory NI, suggesting that classificatory NI can in fact “grammaticalize” into full-blown (object) polysynthesis.

6.2.2. Mithun (1984) on the (non-)syntactic status of NI

Returning to the issue of the syntactic status of NI, for Mithun (1984) NI is “the most nearly syntactic of all morphological processes” (889). First, it “combines constituents (N’s and V’s) usually associated syntactically”, and can be “vastly more productive than other derivational processes since it combines two potentially open sets of morphemes, N and V stems, instead of one set of stems and a limited set of affixes” (889).

On the other hand, according to Mithun, “all morphological processes, however, exhibit a number of special features because of their function as word-formation rules; and NI shares these.” (889). One is that NI is highly productive but not completely free. Second, “the structures created by morphological processes are automatic candidates for

---

Type 2: also has decreasing productivity of compounding, and an increase in opacity of stems;
Type 3: loss of textual frequency, loss of skill in use by bilingual speakers, who rely more on Type 1 and 2 and less on 3 and 4; these may not be acquired at all by children while 1 and 2 may be;
Type 4: new classifiers are no longer introduced, but relics remain in diminishing numbers of sets of classificatory V’s.
lexicalization”, and “speakers know not only whether a derivationally complex word is possible, but whether they have heard it before. The same is not true of unlexicalized syntactic constructions. Few speakers of English are aware that they have never heard a particular sentence before, if the construction, lexical items, and idea are familiar” (889).

Finally, “semantically, lexical items represent conceptual units, unlike unlexicalized syntactic constructions” (889). In other words, NI can often be non-compositional.

These last claims involving “lexicalization” and “morphologization” seem to essentially boil down to the point that morphophonological items can be combined (possibly in a hierarchical structure, although it is not clear that for Mithun even this is necessary) into larger units which are not equivalent to the sum of their parts—i.e. that morphemes can combine to form idioms. Following the pervasively syntactic approach of DM (and much other work in generative grammar), and noting Marantz (1997b)’s claim that even apparently simplex morphemes like cat can be construed as “phrasal idioms”, I will reject Mithun’s arguments that “morphologization” is unbounded by principles of syntax. There is and can be no principled way to distinguish “syntactic” idioms from “lexical” idioms—speakers must know what cat means in the context of the syntactic phrase let the cat out of the bag, which is contextual and therefore Encyclopedic information. Similarly, syntactic Noun-Verb combinations often combine to form “special sound-meaning pairs” which can only be deduced from Encyclopedic information. For example, the phrase close the bar, interpreted compositionally, usually means ‘to establish that the bar is no longer serving by forcing the customers to leave, locking the door, cleaning up, turning out the lights, etc.’, but in the context of a narrator
in a novel by Charles Bukowski, who often writes from the perspective of a hard-core unrepentant alcoholic, it means ‘to sit in the bar, ordering and drinking, until the bartender closes the bar’. The contextual information necessary to correctly interpret such a phrase comes from real-world (i.e. Encyclopedic) knowledge, and as such has no role to play in the underlying syntax of the phrase in question. Thus, it is not the case that “syntax” is intrinsically compositional (i.e. transparent) while “words” are non-compositional, since complex syntactic constructions may have idiomatic interpretations.

Rather than defining word-building in such a way as to preclude syntactic processes, the question seems to be this—to what extent is morphology, including Noun Incorporation, syntactic? It is clear from previous discussion in this dissertation that I take it to be an open hypothesis that all derivation is syntactic, and I think that this question is fundamentally empirical in nature. However, as I noted above, this issue is not limited to NI, but is implicated in other derivational processes as well. If it can be shown that NI is derivational, and that derivation is syntactic, then we may have empirical evidence for the conclusion that NI itself is syntactic.

In section 6.3 I claim that Hopi provides such empirical evidence. However, before we discuss the empirical issues raised by the theoretical questions posed here, I must first demonstrate that NI has identical properties to other derivational processes, specifically, denominal verb formation. This has in fact already been argued by Sadock (1980, 1986), and we will now turn to a review of his direct response to the arguments made by Mithun (1984) (section 6.2.3), as well as Mithun’s rejoinder to Sadock (Mithun 1986) (section
6.2.4), since these earlier discussions will shed further light on the nature of the issues at hand in Uto-Aztecan.


Sadock (1986) gives a direct response to Mithun’s account of NI, reiterating his earlier point, made in Sadock (1980), that the incorporation of nouns to form denominal verbs in Greenlandic is necessarily syntactic. (He also points out that several subsidiary claims of Mithun 1984 were empirically falsified: Greenlandic incorporated nominals can be referential, e.g. they can introduce a new topic in discourse; and that the external syntax of denominal verbs is different from that of non-noun-incorporating verbs—e.g. Greenlandic denominal verbs can be accompanied by elements understood to modify the incorporated nominal, much like the determiners and adjectives modifying possessive denominal verbs that we will see in Hopi below, and throughout Uto-Aztecan in Chapter 7).

According to Sadock, Greenlandic allows incorporation of a direct object into possessive verbs (9), as well as other verbs (10):

(9)  \textit{Qimmeq-arpoq} > qimmeq 'dog'  \\
    'He has a dog' \hspace{1cm} (Sadock 1980: 306 [18])

(10)  \textit{Sapangar-sivoq} > sapangaq 'bead'  \\
    'He bought beads' \hspace{1cm} (Sadock 1980: 306 [19])

When an object noun is incorporated, the sentence is intransitive “with respect to that object, i.e. the verb will never agree with the incorporated object, and thus the object is always understood as indefinite” (Sadock 1980: 307).
There is also an ‘empty stem’ *pi-* which can denote a neutral object, glossed ‘thing’:

(11) \[ \text{Pegarpoq} \]
thing-have-INDIC-3sg
‘He has something’

When a noun is not incorporated and *pi-* is used, the sentence can be transitive, in that a direct object can appear, but the verb is formally intransitive:

(12) \[ \text{Qimmimik} \quad \text{pegarpoq} \]
dog-INST thing-have-INDIC-3sg
‘He has a dog’

That the verb is intransitive is shown by noting that the verb does not have an object agreement suffix, and that the direct object appears with INST-case marking (Sadock 1980).^8

As we will see is the case with the possessive denominal verbs of Hopi, Yaqui, and other Uto-Aztecan languages, the incorporated noun can be modified in Greenlandic. When this occurs, the modifying adjectives take the instrumental case, “the same case that it would have if it were the modifier of a non-incorporated object of a free-standing, formally intransitive verb” (p. 307):

(13)a. \[ \text{Sapanngamik} \quad \text{kusanartumik} \quad \text{pisivoq} \]
bead-INST beautiful-NOM-INST thing-get-INDIC-3sg
‘He bought a beautiful bead.’

b. \[ \text{Kusanartumik} \quad \text{sapanngarsivoq} \]
beautiful-NOM-INST bead-get-INDIC-3sg
‘He bought a beautiful bead.’

---

^8 Greenlandic being an ergative/absolutive language, the direct object of a formally transitive verb normally appears in Absolutive case, as in the following example:

(i) \[ \text{Arnaq} \quad \text{takwara} \]
woman(ABS) see-INDIC.-1sg/3sg
‘I saw the woman’
The grammatical number of the incorporated object is ambiguous without such modification. For these reasons Sadock made the claim that the head noun forms a constituent (NP) with its modifiers in an underlying level of syntax.

In Principles and Parameters-style syntax, this effect is derived by an operation of head movement: the head noun is moved into the verbal position, thus stranding its modifiers (Baker 1988). This is essentially what Jelinek and Escalante (1988) have claimed for Yaqui, and this is the position that I will be defending in reference to (most) denominal verbs in Yaqui and Hopi below. However, this line of reasoning did not convince Mithun that Noun Incorporation itself is syntactic, as we will see in the next section.

6.2.4. Mithun (1986)—Rejoinder to Sadock

Mithun's (1986) rejoinder to Sadock, which Hale and Keyser (2002) cite as a "very credible alternative view" (p. 252) to their syntactic account of denominal verbs (to be discussed in detail in Chapter 8), is essentially that denominal verb formation is a "different formal process" (32) from NI, and as such should not be confused with NI.

Mithun’s distinction essentially boils down to whether the morphological combination involves roots with roots or roots with affixes. Noun incorporation involves the compounding of two identifiable roots, each of which (usually) can appear in isolation (i.e. uncompounded). Denominal verb formation is a process of lexical syntax wherein a nominal root is modified by some (necessarily) bound affix.

Mithun concludes:
In the end, it seems counter-productive to alter the defining features of a structurally and functionally coherent class so that a completely different process can be admitted to that class, and then to abandon the characteristics originally associated with the class because they are no longer exceptionless. Perhaps we can learn more by investigating just why the Greenlandic structure [of denominal verbs] might be, as S[adock] has suggested, so unlike noun incorporation. (p. 36, emphasis added)

However, it seems to me that this same argument could be applied to Mithun’s very notion of “noun incorporation”, which for Mithun includes lexical compounding, which she herself shows is not limited to nouns (e.g. in N + V compounds) but also includes other categories as well: V+V, V+A, etc. Are these to be understood in terms of “incorporation” as well? If not, why not? In conflating compounding with “noun incorporation”, Mithun has taken what she assumes to be a lexical process (compounding) and has defined what may or may not be lexical (incorporation) in such a way as to presuppose that it is and can only be lexical. That is, Mithun’s definition of incorporation as “everything that is lexical” about the merging of heads as “incorporation” and “anything that is syntactic” about the merging of heads as “something else” that doesn’t fit the definition, and which should therefore be considered a peripheral construction, is circular.

Furthermore, from the PSP, the distinction between roots and affixes, while important in discussing the morphophonological appearance of morphemes, is not sufficient to justify a Lexicalist view of syntax. Mithun (1999), following Carlson (1990), shows
persuasively that "lexical affixes" have their origins in earlier stages of compounding among lexical roots. Thus, affix-hood is itself only one stage in a diachronic continuum of root-ness and affix-ness. From a view of synchronic syntax that denies the "Lexicon" as a locus of syntax-like operations before syntax proper, one may ask to what extent the root/affix distinction, and therefore the compounding/affixation distinction, is actually necessary in teasing apart the (grammatical) location of morphological processes. In fact, treating this distinction as irrelevant in synchronic analyses is more consistent with the diachronic view propounded in grammaticalization theory, since Mithun's distinction requires a discrete discontinuity between these categories—one day a process (e.g. noun incorporation) is morphological and the next day, when the affix is bound (and is therefore a derivational affix rather than a free-standing root), it is not. No such discrete break is necessary from the syntactic perspective being advocated here.

Presupposing for a moment that Mithun would grant that denominal verb formation is (or at least can be) syntactic in a non-Lexicalist sense (i.e. thus granting Sadock's point regarding the syntactic character of denominal verbs in Greenlandic), what would seem to be needed to show that NI is also (or at least can be) syntactic would be a language that has identical properties for both classes of word-formation: noun incorporation and denominal verb formation. If this could be done then Mithun's identification of the two classes as separate ontological categories would be called into question, and the possibility (not to mention the desirability) of a unified analysis would be greatly enhanced.
In fact, in section 6.3 I will show that there is such a language, Hopi (NUA) (data from K. Hill 2003), and I argue that the best analysis of incorporating and denominal verbs is the one that unifies the two. I argue that a syntactic account is the most perspicuous way to derive this unification. Although Sadock makes a similar point regarding Greenlandic, from the framework of Autolexical Syntax, following the PSP (e.g. Baker 1988, Hale and Keyser 1993, 2002, and other work) I promote a head-movement account within DM, and I will show that this account predicts the kind of analogous constructions that we see in noun incorporation and denominal verbs in a language like Hopi. In Chapter 7 I examine the extant empirical evidence to see how far this approach can take us in examining the noun incorporating and denominal verbs of the Uto-Aztecan languages. In that chapter I will argue that there are (at least) four observationally distinct types of NI in Uto-Aztecan, which have somewhat different properties than the more broadly cross-linguistic categories given by Mithun (1984). The formal theoretical machinery necessary to derive the distinct NI types that we find in Uto-Aztecan is the topic of Chapter 8.

My ultimate conclusion is that we must recognize two formally distinct types of NI—those of the head movement kind, which I will continue to call incorporation (following Baker 1988, Hale and Keyser 1993, 2002), and those of “classical noun incorporation”, as identified by Sapir (1911), which involve compounding. I distinguish this type of NI from the former by positing an analysis involving the merging of lexical roots for this kind of NI, which I will call conflation (following Hale and Keyser 2002). Thus, the two
types of NI are distinguished by the two processes involved—Move (incorporation) or Merge (conflation).[^9]

I should note here before proceeding that these terms have presented a bit of a quandary, in that the term incorporation has been used by various linguists to denote processes which overlap both movement and compounding. In so far as the term incorporation implies some kind of operation (perhaps taking one thing from one place and putting it somewhere else!—i.e. movement, although this is certainly not what Kroeber had in mind in 1911 when he referred to incorporation as a “process”), I will use this term to refer to moving a head from one phrase into the head of another, in the technical sense of Baker (1988).

To capture the sense of “incorporation” as compounding held by Mithun, on the other hand, I will apply the Minimalist Program-specific term conflation, which I take to be the technical implementation of compounding in sentential syntax, via Merge. Whereas compounding is generally understood to involve the merger of two or more identifiable roots to form a single constituent, conflation refers to the merger of roots to (possibly phonologically null) heads, to form complex heads. This view of compounding is theory-specific, but seems to be consistent with the general understanding of what compounding is. What may be controversial here is my stance, possibly empirically falsifiable, that this

[^9]: Technically, Move may not be a distinct process from Merge. Under Chomsky (2001)’s conceptualization, there is a distinction between external and internal set merge: where the former combines two independent objects, and in the latter, “if β is part of α. . . β is said to be a ‘copy’ of its occurrence in α” (Langendoen 2003: 308). Internal set merge is essentially the technical implementation of “Move” within set theory.
compounding takes place in syntax proper and not in some separate generative "Lexicon".

Before embarking on further discussion of the theoretical proposals being advocated to account for noun incorporating and denominal verbs, however, we must first examine the data that suggests that they are in fact formed by the same process, as has recently been suggested (e.g. by Hale and Keyser 1993), in order to refute Mithun's bifurcation of the two processes into separate ontological categories.

6.3. Denominal and noun incorporating verbs in Hopi: Evidence for a unified analysis

K. Hill (2003), like Sadock (1980, 1986) and contra Mithun (1984, 1986), regards Hopi denominal verb formation as a process similar (if not identical) to noun incorporation, and his article on the topic in Hopi is the most thorough discussion of the similarity of these two processes available for any Uto-Aztecan language.

Like polysynthetic languages, "Hopi allows for incorporation of nouns into verb structures, but it lacks corresponding morphological structures with pronominal elements" (p. 218), i.e. Hopi lacks pronominal arguments. A polysynthetic language typically allows for either a noun root or a pronominal affix to be the patient argument attached to any transitive verb, as in the following Nahuatl examples:

10 There is, however, a small exceptional class of about five "pronominal verbs" which take pronominal prefixes (K. Hill 2003: 221-222). These will be discussed with respect to the development of polysynthesis in Nahuatl in Chapter 9.
(14)a. *Ni-c-tequi.*
   I-it-cut
   I cut it.

   b. *Ni-xōchial-tequi.*
      I-fruit-cut
      I pick fruit.

See further discussion of the polysynthetic properties of Nahuatl in Chapter 9.¹¹

An example of a Hopi NI construction is given in (15b); (15a) shows that neither an incorporated nominal nor pronominal argument marking for direct objects are obligatory for transitive verbs in Hopi:

(15)a. *Nu’ pu-t tsakwna.* (K. Hill 2003: 218 [12])
   I that-ACC dry
   I dried it out.

   b. Askwali, nu’ wuhuaq sipsl-tsaakwna.
      thank.you(FS), I many peach-dry
      Thank goodness, I’ve dried out a lot of peaches (for storage). > sipala ‘peach(es)’

Since pronominal marking (or object incorporation) is not required on a transitive verb, Hopi is not polysynthetic in the sense of Baker (1996).

Hill makes a distinction between denominal verbs and noun-incorporating verbs, although he acknowledges that the two “have much the same syntactic properties” (225). His distinction seems to be that whereas NI verbs involve incorporation into a lexical verb stem, denominal verbs involve verbal inflectional elements affixed directly to the

¹¹ Gronemeyer (1996) attempts to regard Hopi as a polysynthetic language, apparently in order to maintain Mithun’s implicational hierarchy of NI types, since Hopi has Type 4 (classificatory) NI. However, as I will show in detail in Chapter 7, other non-polysynthetic Uto-Aztecan languages also have classificatory NI without being polysynthetic, thus refuting Mithun’s implicational hierarchy, and also Gronemeyer’s claim that “syntactic NI only occurs in polysynthetic languages” (37).
noun root—i.e. Hill’s distinction appears to be the same as that made by Sapir (1911) and Mithun (1986), among many others.

I make the theoretical claim that this distinction is illusory, and that Hill’s intuition regarding the closeness of the two with regard to their syntax is based on the fact that they are both formed in sentential syntax through noun incorporation in the head-movement sense—i.e. these patterns of denominal verb formation are examples of syntactic noun incorporation. Let us review the properties of each of these constructions in order to see clearly their similarities, in order to understand why I would want to claim that they are in fact the same.

6.3.1. Noun incorporation in Hopi

Hopi incorporated nouns are usually the direct object of a transitive verb, as in (16):

(16)a. *humit-tuvahoma*
    shelled.corn-wash
    ‘wash corn kernels’
    > *humita* ‘shelled corn’

b. *kwew-yuku*
    belt-finish
    ‘finish making a belt or belts’
    > *kweewa* ‘belt’

c. *kaway-’u’uyi*
    horse-steal
    ‘steal a horse or horses’
    > *kawayo* ‘horse’

d. *tap-nina*
    cottontail-kill(SG.OBJ)
    ‘kill a cottontail’
    > *taavo* ‘cottontail’

e. *tap-qöya*
    cottontail-kill(PL.OBJ)
    ‘kill cottontails’
    > *taavo* ‘cottontail’
According to Hill, for ditransitive verbs, “it appears that only the direct object can be incorporated” (K. Hill 2003: 227, emphasis added):

(17)a. Ung mupi-mqa.  
     you(ACC) rolled.up.thing-give  
     ‘She gave you a roll of piki.’  
     > mûpi ‘s.th. rolled up, e.g., piki’

b. Nu’ i-tiw’aymu-y sòosokm-u-y siva-huyta.  
     I my-maternal.siblings-ACC all.of.them-ACC money-distribute  
     ‘I gave money to all of my nieces and nephews.’  
     > siiva ‘metal, silver; money’

The head noun of NI constructions may be modified, and the modifiers appear in accusative case (when possible):

(18) Naat itam pu-t qatsi-yese.  
     still we that-ACC life-sit(PL)  
     ‘We are still living that life.’  
     > qatsi ‘life’  
     > yeese ‘sit, live (PL)’

(19) Nu’ pay i-t tôövu-t aw qötö-tpe.  
     I well this-ACC embers-ACC to.it head-roast  
     ‘I roasted this head over the embers.’  
     > qötö’at ‘3p’s head’

I follow Hale and Keyser (2002) in assuming that this is evidence that the head noun originates in direct object position, and incorporates (via head movement) into verbal position.

Finally, Hopi displays classificatory noun incorporation, in that a verb formed via NI may take an overt direct object nominal direct object. In these cases, a “a relatively general [noun] stem is incorporated to narrow the scope of the [verb]. . . but the compound [noun + verb] stem can be accompanied by a more specific external NP which identifies the argument implied by the [incorporated noun]” (Mithun 1984: 863). An example is given in (20):

(20) Nu’ yöypala-t kuy-tàngta.  
     I rainwater-ACC contained.liquid-put.into.container(s)  
     ‘I put the rainwater into some containers.’  
     > kuyyí ‘contained liquid’
Such facts as this kind of classificatory NI would seem to negate the possibility of a head movement account, since we do not see root-identity between the incorporated nominal and the over nominal in the direct object position. However, I present an analysis in Chapter 8 that reconciles this problem, which I term the *hyponomous argument problem*, which is equally relevant to denominal verbs.

In the next section I show that similar facts obtain with respect to denominal verb constructions, leading to the conclusion that denominal verbs are formed in the same way as NI verbs.

6.3.2. Denominal verbs in Hopi

Hill (2003) provides a variety of forms of “denominal morphology”, including the following for non-possessed nouns (2003: 225-6 [39-46]):

<table>
<thead>
<tr>
<th>(21)</th>
<th>perfective</th>
<th>imperfective</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ta (pl. -tota) ‘causative’</td>
<td>-lawu (pl. -lalwa) ‘continuative’</td>
<td></td>
</tr>
<tr>
<td>-ti (pl. -toti) ‘realized’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(22)</th>
<th>kii-hu ‘house’</th>
</tr>
</thead>
<tbody>
<tr>
<td>kii-ta ‘build a house’</td>
<td></td>
</tr>
<tr>
<td>kii-ti ‘become a house’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(23)</th>
<th>qöötot ‘suds, lather, foam’</th>
</tr>
</thead>
<tbody>
<tr>
<td>qöt-ta ‘make suds, lather up’</td>
<td></td>
</tr>
<tr>
<td>qöt-ti ‘become sudsy, foamy’</td>
<td></td>
</tr>
</tbody>
</table>

Possessed nouns, on the other hand, are followed by the ‘possessed’ suffix -'y-:

<table>
<thead>
<tr>
<th>(24)</th>
<th>-(i)va (pl. -(i)va-ya) ‘ingressive’</th>
</tr>
</thead>
<tbody>
<tr>
<td>-(i)ta (pl. -(i)yungwa) ‘durative’</td>
<td></td>
</tr>
<tr>
<td>-(i)ma (pl. -(i)wisa) ‘progressive’</td>
<td></td>
</tr>
<tr>
<td>-(i)numa (pl. -ya) ‘circumgressive’</td>
<td></td>
</tr>
</tbody>
</table>
(25) awta-‘at ‘his bow’
awta-‘y-wa ‘arrive with a bow; get to have a bow’
awta-‘y-ta ‘have (as) a bow’
awta-‘y-ma ‘go along with a bow’
awta-‘y-numa ‘go around with a bow’

Hale and Keyser (2002) analyze the -‘y suffix as the possessive verbalizer, and I follow this analysis here.

Another pattern of denominal verb formation can be seen with two homophonous (in the singular) suffixes, which Hill calls “perfective –ta”, with plural –tota and glossed ‘causative’, and “imperfective –ta”, with plural -yungwa and glossed ‘durative’. When non-plural these morphemes can be disambiguated by a difference in tone, where the perfective has a falling tone and the imperfective does not (236):

(26)a. kweewa ‘belt’
   b. kwew-ta (pl. -tota) ‘put on as a belt, make into a belt’
   c. kwew-ta (pl. -yungwa) ‘have on as a belt, have something around the waist’
   d. cf. kweewa-ta (pl. -tota) ‘make a belt’

Other derivational processes can further affect the valence of the verb. For instance, applicatives can be derived by means of the addition of another causative suffix, -ina, which when combined with the suffix -ta results in a complex suffix -toyna (pl. ~ya) (236):

(27) causative -ta (pl. ~tota)
   a. kiita ‘build a house’
   b. kiitoyna ‘build a house for, provide with a house’
   c. ãya ãayata ‘make a rattle’
   d. aayatoyna ‘provide with a rattle’

(28) repetitive -ta (pl. ~tota)
   a. hanta ‘take down (pl.obj.)’
   b. hantoyna ‘be taking down (pl.obj.)’

12 It is interesting to note that the form of the root kweewa changes with some affixes but not others.
(29) durative -ta (pl. -yungwa)
   a. ngu 'yta ‘be holding (something)’
   b. ngu 'ytoyna ‘be holding (something) for (someone)’

We will see in Chapter 7 that some of this morphology goes a long way back in Uto-Aztecan, since these morphemes have cognates in many other Uto-Aztecan languages.

As is the case with NI in Hopi, the head noun of denominal verbs may be modified by adjectives, quantifiers, etc. These appear in ACC case here as well:

(30) Hak i-t kis-ta? (K. Hill 2003: 234 [96])
   who this-ACC shade-CAUS ‘Who built this shade?’ > kiisi ‘shade’

(31) Um qa hii-ta ho-'y-va? (K. Hill 2003: 234 [97])
   you not what-ACC arrow-POSS-INGR ‘Didn’t you bring any arrows?’ > hoohu ‘arrow’

In addition, denominal verbs also appear in a construction reminiscent of classificatory NI, where the head noun in the direct object noun phrase indicates a subset of the nominal root “incorporated” into the denominal verb:

(32) Pam tsiili-t nakwa-'y-ta-ngwu. (K. Hill 2003: 236 [112])
   he chile-ACC feather.worn.on.head-POSS-DUR-HAB ‘He (the Hehey’a kachina) wears chili pepper (“chile as a feather”) on his head.’ > nakwa ‘feather worn on the head’

(33) Nitkyamokpi-v kwèw-ta. (K. Hill 2003: 237 [116])
   bundled.journey.food-ACC belt-CAUS ‘He fastened his bundled journey food around the waist.’
   (“He put on his bundled journey food as a belt.” OR “He made his bundled journey food into a belt.”) > kweewa ‘belt’

Thus, denominal verbs have the same syntactic distribution as denominal verbs, indicating a similar process.

In sum, like Greenlandic, Hopi allows for external modification (i.e. “stranding”) for incorporated nouns in NI constructions and in denominal verb constructions. Unlike
Greenlandic, however, these modifiers appear in Accusative case in Hopi, indicating that the verb is transitive, unlike Greenlandic where the stranded modifiers appear in an oblique (instrumental) case, which led Sadock to the conclusion that such verbs are intransitive in Greenlandic.

6.3.3. Hopi denominal verbs as incorporating verbs

Although there is probably some (at least subtle) difference between the incorporated and non-incorporated constructions with respect to their use in discourse, according to K. Hill (2003), "whether an object is a separate word or is incorporated into the verb seems to make no difference" (238), cf. (34) and (35):

(34) Pam nuy suta-t qōma-to-yna. (K. Hill 2003: 237 [127])
    he me red.ocher-ACC face.paint-CAUS-CAUS
    'He applied red.ocher to my face.'  > qōma-'at '3p's face paint, makeup'

    yet he our-older.brother-ACC red.ocher-face.paint-RDP-CAUS-CAUS
    'He is still applying red.ocher to our older brother's face.'

Unlike Yaqui, which has a verb, hippue, indicating 'have', Hopi does not have a verb indicating possession, and requires instead the incorporation of the head of the object DP onto the affix -y- (discussed above). Thus, incorporation is obligatory with the possessive morpheme -y-. In Chapter 8 I claim that this obligatory incorporation is a lexical property of the possessive morpheme, characterized by Hale and Keyser (2002) as having a defective “phonological signature”. This property induces obligatory incorporation, and is in effect a stipulation that makes the morpheme an affix. Other Uto-

---

13 Hopi is not alone in Uto-Aztecan in this regard. Nahuatl did not have a verb have until contact with Spanish, when suffixal -piya was calqued on Spanish tener (Jane H. Hill, personal communication).
Aztecanists have also regarded such morphemes in other Uto-Aztecan languages as obligatorily incorporating verbs, such as Dayley (1989) for Tümpisa Shoshone. However, this property does not in any way necessitate the separation of a lexical morphology module from syntax proper, as Mithun’s root/affix distinction would otherwise have us do.

Returning to the properties of incorporation in Hopi, Hill also shows that these incorporated nominals can have definite reference, which is cross-linguistically atypical (but also possible in Huauhtla Nahuatl, Merlan 1976, and also Greenlandic, Sadock 1980):

(36) \[ Nu' \quad \text{pakiw-maquito-ni;} \quad \text{(K. Hill 2003: 241 [143])} \]
I \quad \text{fish-go.hunting-FUT;}

\[ \text{noqw itam pu-t enang nöönösa-ni.} \]
so we that-ACC in.addition.to eat(PL)-FUT

‘I’m going fishing, so we can eat it (fish) along with the other food.’
> paakiw 'fish'

All of the facts discussed here for Hopi lead to the desirability of a unified analysis of the Noun Incorporation and denominal verb constructions. Such an account is suggested by Hale and Keyser (1993, 2002), and I will argue in Chapter 8 that the classificatory and hyponomous argument data are compatible with an analysis utilizing syntactic head movement—i.e. incorporation.

We turn now to a discussion of one final theoretical contribution to the literature on noun incorporation—the Lexicalist analysis of Rosen (1989).

Having argued thus far in this chapter that noun incorporation and denominal verb formation are formed by identical processes, and claiming further that these processes are syntactic, it is now time to consider what the most perspicuous syntactic account of these phenomena may be. Since the publication of Chomsky (1970), the two major competing schools of thought concerning word-formation in generative grammar have been those of the “Lexicalists”, who hold that morphological items are formed in the lexicon to be inserted into syntactic terminal nodes, and those who hold the more restrictive view of the Pervasive Syntax Perspective, arguing that there is no combinatory mechanism outside of syntax proper.

In this dissertation I have adopted the latter position, working within the Distributed Morphology framework, and I will develop proposals for Uto-Aztecan where NI and denominal verb are formed in the syntax. However, I will first discuss a prominent Lexicalist account of NI, Rosen (1989), and will argue that the pervasive syntax approach is to be preferred.

Rosen (1989) divides cross-linguistic NI into two fundamentally different types: Compound NI and Classifier NI. In Compound NI, so-called because of its similarity to compounding, the N(oun) + V(erb) complex is intransitive, as the process of NI affects the argument structure of the verb (i.e. the incorporated noun acts as an argument of the verb). In Classifier NI, on the other hand, the transitivity of the verb is not affected, and the complex verb can co-occur with a direct object argument. When this happens “the direct object argument and the incorporated noun are linked semantically in much the
same way that a noun classifier is linked semantically to the noun it classifies” (296).

That is, the meaning of the direct object is (usually) semantically hyponomous with respect to the meaning of the incorporated noun. Compound NI is correlated with Mithun (1984)’s Types 1-3 NI, and Classifier NI with Mithun’s Type 4 NI (see discussion in section 6.2.1). Whereas Rosen bases her distinction entirely on the syntactic properties of these two classes, she argues that Mithun’s previous discussion adds further functional and discourse-based support for the Lexicalist analysis of NI.

As Rosen notes, languages with Classifier NI exhibit “stranding”, which she defines as “a process whereby an NP modifier is left with no head noun, while a noun with the same semantic reference is incorporated into the verb” (296). Rosen calls these “null head modifiers”, and she emphasizes that these null head modifiers are independently attested in these languages so they must be possible with or without noun incorporation.

This is certainly the case for Yaqui, as the following data illustrate:

(37)a. \textit{aapo 'uka vicha-k} \\
3sg DET-ACC see-PERF \\
‘He saw that [one]’.

b. \textit{aapo siali-k vicha-k} \\
3sg green-ACC see-PERF \\
‘He saw a green [one]’.

Similar examples were given for Hopi above, in (15)a and b. One piece of evidence that Rosen uses to argue for a “null head modifier” analysis is that adjectives in Mohawk show gender agreement with non-existing nouns. (However, one wonders if there are default features, or if features could be inferred from discourse context, in these cases.)
Rosen’s Lexicalist theory “attributes the stranding facts to an independent phenomenon in the language (the existence of null head modifiers in general)” (298). According to Rosen’s view,

[i]f one were to subscribe to a syntactic account of NI, stranding of modifiers in an NI construction and the independent existence of null-headed NPs would require two different accounts. In a syntactic account, the stranded modifiers are created by the movement of the head out of the NP, thereby leaving modifiers behind. Stranding is then directly tied to syntactically derived NI. However, in the lexical theory proposed here, null-headed modifiers in NI constructions and those independent of NI constructions both have the same source. (p. 301)

Another issue raised by Rosen for such languages is that of transitivity. According to Rosen, “there is no indication in the literature that incorporation ever affects the transitivity of the verb in these languages—a transitive verb remains transitive after NI has applied, and an intransitive verb remains intransitive after NI has applied” (p. 302). For Rosen, languages fall into classes as to which kind of NI they have—Classifier NI or Compound NI. Rosen does not seem to address Mithun (1984)’s claim that Classifier (or “Type 4”) NI entails Compound (Types 1-3) NI. However, we will see in Chapter 8 that the Uto-Aztecan languages provide counterevidence to Mithun’s proposed entailment hierarchy, since non-polysynthetic Uto-Aztecan languages (e.g. Hopi, the Numic languages, and others) have Classifier NI, and therefore polysynthesis is not a prerequisite for Classifier (Type 4) NI. Nevertheless, it is also not the case that there is complementary distribution of these two NI types, as Rosen suggests, since some Uto-
Aztecan languages (at least Yaqui) have both. One benefit of the syntactic account that I develop below is that the appearance of both NI types is attributable to universal properties of syntax, parameterized in particular ways in particular languages, rather than positing different “language types” which might lead one to suspect a necessary separation of the NI types.

The final characteristic of Classifier NI is “doubling”, where the incorporated verb takes a direct object and thus has no source for syntactic incorporation. That is, doubling “gives further evidence for the syntactic independence of NI and the direct object NP position; it is an example of NI without a null element” (302). In addition, according to Rosen, examples of doubling “highlight the lexical nature of Classifier NI: there is no syntactic source for such doubled nouns” (303).

This issue, which I call the *hyponomous argument problem*, certainly raises a serious issue for the syntactic account of NI, and the identical critique can be raised for the incorporation account of the derivation of denominal verbs. Hale and Keyser (2002) separate this process from the *cognate argument problem*, where there is root identity between an incorporated noun (or denominal verb) and its direct object. They suggest that an analysis could be posited in which the latter is formed by syntactic movement. In the former case, however, they feel that movement cannot derive the incorporated noun—e.g. in examples like *John danced a jig*, where the denominal verb *dance* had been previously analyzed as noun incorporation from the underlying object position (Hale and Keyser 1993), Hale and Keyser abandon the noun incorporation account in their 2002 monograph. I will suggest in Chapter 8, however, that this abandonment of the movement
account was premature, and I will offer a Late-Insertion theory of NI and denominal verbs in hyponomous argument constructions.

What is crucial in the discussion of Rosen's NI types here, however, is to disentangle the issues of where these operations occur (for me, it is in syntax; for Rosen, in the Lexicon) and from what the relevant syntactic operations are in the first place. The Uto-Aztecan languages, and especially Hopi, provide interesting test cases to apply the PSP approach to data that call into question the distinctions made between incorporation, compounding, and derivational morphology in earlier theories.

A further problem raised by Rosen is that there is "no syntactic source for the semantic restriction" (304) on selection of the object in NI—i.e. why must a direct object NP (or DP) be a subset of the noun incorporated into V? In a syntactic account, where the semantics are read off of the syntax, this might follow directly, if there is some universal restriction that verbs doubly modified in this way force a hyponomous (or equivalent) interpretation when a nominal root is merged with the verb and a full DP appears as sister to this complex verb. On the other hand, it is not as easily explained in a Lexicalist account, where verbs might be expected to take whatever semantic meanings that they "feel like", so to speak, rather than being constrained by some formal properties internal to the language faculty itself (i.e. the universal syntactic structure being sought by practitioners of Principles and Parameters-style syntax).

Rosen is forced to admit that there are apparent counterexamples to her Lexicalist approach. Regarding two of these, she says:
Southern Tiwa and West Greenlandic diverge from the canonical Classifier NI languages on the property of doubling: they do not allow any doubling at all. This poses an apparent counterargument to the analysis of Classifier NI presented here. Any language in which the verb retains its transitivity and allows stranding with NI is predicted to have doubling as well. (p. 306)

Rosen considers the possibility that NI in these languages may in fact be due to syntactic head movement a la Baker (1988) (or co-analysis in Autolexical Syntax, a la Sadock 1985, 1991), although she highlights an issue that is raised in such a theory, insofar as such a theory must also allow for doubling in other languages as well: “The properties that [syntactic incorporation] would have to explain are (i) the fact that NI itself does not affect transitivity of the verb; (ii) the stranding facts; (iii) the lack of doubling; and (iv) the specific selectional restrictions that the incorporated noun places on the verb” (308-9).

In Compound NI, on the other hand, “the direct object argument of the simple verb is satisfied, so that no direct object can co-occur with NI” (309). These verbs should therefore always be intransitive. Further, “in languages with Compound NI there should be no stranding of determiner or of modifiers” (311), which appears to be true for the Polynesian languages.

However, this claim presents a problem for Rosen: Yaqui has exactly this kind of incorporation, where no modifiers are allowed and a normally transitive verb appears with intransitive morphology. Consider the data in (38):
(38) a. aapo maaso-ta peu-ta-k  
   3sg deer-ACC butcher-TRAN-PERF  
   'He butchered a deer'

b. aapo maaso-peu-te-n  
   3sg deer-butcher-INTR-PAST  
   'He was deer butchering'

c. *aapo bwe’uu-k maaso-peu-te-n  
   3sg big-ACC deer-butcher-INTR-PAST  
   ["'He was [big deer]-butchering"] or ["'He was deer-butchering a big one’]

The patterns exhibited by the Yaqui data in (38) are classical root-root NI, and the verbal morphology indicates the change in transitivity that occurs under NI. When the noun root is incorporated into the verb the verb is intransitive, and is marked with the intransitive –te suffix (rather than transitive –ta) (Jelinek 1998). However, it is the case that Yaqui also has Classifier NI, and allows for stranded modifiers (even null head modifiers for normal transitive verbs, as in 37).

Rosen suggests that

*if* a Compound NI language did have null-head modifiers, there is a clear prediction concerning the interaction of NI and null-head modifiers: they would not interact at all. Because the argument is satisfied by the NI word-formation process, there is no NP related to the incorporated noun, and thus there can be no NP with a null head related to the incorporated noun (312, emphasis added)

According to Rosen, neither Baker (1988) nor Sadock (1985) capture the “cluster of grammatical properties” Rosen thinks she has shown to be related to NI. These are summarized in Table 6.1:
I grant Rosen’s points regarding Compound NI—as she suggests, and as has been believed since at least Sapir (1911), this kind of NI is essentially indistinguishable from lexical compounding. We simply differ on where this compounding takes place. For Rosen it occurs in the Lexicon, before syntax. In DM, there is nowhere for it to occur except in sentential syntax.

It is not my suggestion here that a Lexicalist approach cannot account for compound and classifier NI. However, it appears to me that what we need is a unified explanation, which I propose is the syntactic one. It seems to be the case that what Rosen has done in describing her two NI types is re-state the facts: compound NI, by definition, is root-root (N+V) compounding which, by definition, leaves no room for modifiers. The second crucial claim, however, is that hyponomous arguments cannot be derived via syntactic movement. I will make the novel theoretical claim in Chapter 8 that this is in fact possible. Thus, the two sources of complex head formation posited in the recent Minimalist literature, Merge and Move, are all that are required to derive the two classes of NI described by Rosen.

As to the Classifier NI, it is my claim that the “cluster of grammatical properties” that Rosen has identified, in effect, are the result of a single fact, rather than three. This is

<table>
<thead>
<tr>
<th>Classifier NI</th>
<th>Compound NI</th>
</tr>
</thead>
<tbody>
<tr>
<td>*argument structure unaffected by NI</td>
<td>*argument satisfaction within the complex verb;</td>
</tr>
<tr>
<td>*stranding of modifiers</td>
<td>*no stranding of modifiers</td>
</tr>
<tr>
<td>*doubling outside of verb</td>
<td>*no doubling outside of verb</td>
</tr>
</tbody>
</table>

*Table 6.1: Cluster of properties associated with noun incorporation* (Rosen 1989: 313)
that head movement has taken place in syntax. The argument structure of the verb is "unaffected" because, by definition, a transitive verb has a direct object (NP). When the head N incorporates into V, it "strands" whatever modifiers appear in NP (or DP) with it. As for the issue of doubling, in the case of hyponomous arguments, where there is true classifier NI, the fact that any non-cognate argument that appears in the "trace" position of the incorporated noun is hyponomous with respect to that noun will be derived via the notion of Late-Insertion: the direct object argument must be a more specific sub-set of the class of entities denoted by the incorporated noun, because it shares at least the same features as the incorporated noun. All of the features of the incorporated noun will be shared with the direct object noun (or NP), which can have further feature specifications to make it more specific (see Chapter 8).

In sum, my claim is that, assuming the PSP, all we need to do is identify the operations used to derive the NI types identified by Mithun and Rosen. What I will do in Chapter 8 is argue that two notions, Merge and Move, are all that we need to derive the range of facts that have been discussed by Mithun and Rosen. My approach also allows for an explanation of the additional facts that the Lexicalist approaches miss: e.g. the "doubling"-free stranding facts of Southern Tiwa and West Greenlandic. Thus, following the neo-Kroeberian program of universalizing the structures and operations available in syntax (i.e. UG), I will argue that all of all of this is done in syntax proper and that the notion of a Lexical Syntax is completely superfluous.

Before moving on, I would like to make a brief note on the concept of "Lexical Syntax". Hale and Keyser (1993) present a theory of argument structure which relies on
a domain of syntax that has come to be identified as "L-syntax" (i.e. "Lexical Syntax")—see discussion of this notion in section 8.2.1 below. One aspect of this notion that was crucial for Hale and Keyser (1993) was that this domain was regulated by the same principles that operate in syntax proper, e.g. the Head Movement Constraint (Travis 1984). Other views of the Lexicon, and Lexical Syntax, exist. For Mithun and Corbett (1999), for example, the Lexicon is simply the set of "lexical items" and rules for creating lexical items, the latter of which "differ crucially from syntactic processes in that they are not the primary mechanism for producing on-line speech" (68). Rather, "they create lexical items to be learned, stored, and accessed as units" (68). Since DM is a model of competence, rather than performance, it is not clear to what extent the criteria laid forth by Mithun and Corbett are relevant to the claims to be made below.

6.5. Summary

To summarize the discussion so far, from a Sapirian perspective Mithun (1986) has made an epistemological argument that "noun incorporation" and "denominal verb formation" are intrinsically different kinds of morphological processes, and should not be confused or conflated with one another. According to this view, previous arguments (e.g. Sadock 1980, 1986) that have attempted to link the two have made a category mistake. According to the data presented by K. Hill (2003), on the other hand, Hopi displays identical

14 In concluding their 1993 work, Hale and Keyser noted that it would be ideal to abandon the separate notion of L-syntax and have all syntax operate at the same level, and by their 2002 publication they had undertaken this step.

15 However, Pfau (2000) discusses DM as a model of production, and gives psycholinguistic evidence that regular inflectional and derivational morphology is computed on-line.
characteristics between the two processes and thus strengthens the basis for approaching these data as a unitary phenomenon, since they are apparently not as distinct as Mithun has claimed.

In Chapter 8 we will review recent discussions that have focused on the syntax of noun incorporation, some of which have explicitly linked this to denominal verb formation. However, before discussing these theoretical issues, in Chapter 7 I take a comparative perspective in order to show that the similarities among N1 and denominal verb constructions are not limited to Hopi, but are in fact typical of the Uto-Aztecan languages. This leads to the reconstruction of certain syntactic constructions for Proto-Uto-Aztecan.
CHAPTER 7
NOUN INCORPORATION IN UTO-AZTECAN

7.1. Introduction

In the last chapter I surveyed two prominent proposals for categorizing cross-linguistic noun incorporation (NI) into distinct sub-types. The first of these are the discourse-functional categories proposed by Mithun (1984): Type 1 NI, equivalent to Sapir’s notion of ‘noun incorporation’ and involving noun-verb compounding; Type 2, involving the manipulation of case relations; Type 3, involving the manipulation of discourse structure and linked with polysynthesis; and Type 4, “classificatory NI”. Mithun’s examples of classificatory NI involve incorporation of actual classifiers (i.e. morphologically obligatory elements that specify a range of possible objects), but I argue that “classificatory NI” is also an appropriate term for NI with hyponomous objects—i.e. NP objects that denote some subset of the class of items denoted by an incorporated noun. As discussed in Chapter 6, Mithun made the strong claim that these NI types form an implicational hierarchy: Type 1 >> Type 2 >> Type 3 >> Type 4.

Mithun’s discourse-functional categories have subsequently been addressed by Rosen (1989), who divides NI types into two kinds based solely on distributional and syntactic properties. Rosen’s categories are Compound NI, which changes the valence of verb (i.e. intransitivizes it), and Classifier NI, which does not change the valence of the verb, thus allowing for stranded modifiers (“stranding”) and cognate and hyponomous arguments (“doubling”).
In this chapter I present my own proposals for categorizing NI into types relevant to the Uto-Aztecan language family. I define and illustrate these in section 7.2. In section 7.3 I will draw some preliminary conclusions regarding the cross-Uto-Aztecan NI facts. The data collected in my survey of UA NI are presented and discussed in section 7.4.

7.2. Categories of Noun Incorporation in Uto-Aztecan: Overview

Like Rosen (1989), I divide NI types into categories based on their distributional properties, thus I view NI types in syntactic terms. These NI types roughly parallel the discourse-functional NI types proposed in Mithun (1984), although some differences do arise. Although these NI types are defined according to whether or not they meet certain distributional criteria, they are offered here as heuristic categories only. In Chapter 8 I suggest that a Rosen-style bipartite categorization is in fact correct, although I will present a non-Lexicalist syntactic analysis of these phenomena. For descriptive purposes in linguistic comparison, however, I have found the following categories to be of use, and it is conceivable that there may be more that could also be considered, at least in other language families.

The categories are N-V compounding (7.2.1), syntactic NI (7.2.2), “object polysynthesis” (7.2.3), and classificatory NI (7.2.4).

7.2.1. N-V Compounding

The first type of NI relevant to Uto-Aztecan is what I call N-V compounding. This term corresponds almost exactly to the canonical notion of noun incorporation found in Sapir
(1911), and is almost equivalent to Mithun (1984)'s "Type 1 NI" and Rosen's "Compound NI". Although I use the term *compounding* NI in contrast to *syntactic* NI (discussed in section 7.2.2), this should not be taken to indicate that I view compounding as somehow "lexical", in the sense of "pre-syntactic", as others have argued (e.g. Rosen 1989, Mithun and Corbett 1999). I will delay my syntactic analysis of these phenomena until Chapter 8. In the meantime, this categorization is of descriptive use and does not depend on any theoretical assumptions as to where the compounding takes place with respect to other morphosyntactic processes (i.e. whether it is lexical or non-lexical).

N-V compounding can come in at least two varieties: object incorporation and instrument/manner incorporation. In the first of these, the noun is incorporated into the verb, and the verb is formally intransitive. For example, consider the Yaqui data below, repeated from the previous chapter:

(1) N-V Compounding in Yaqui

a.  *aapo maaso-ta peu-ta-k*  
   3sg deer-ACC butcher-TRAN-PERF  
   'He butchered a deer'  

b.  *aapo maaso-peu-te-n*  
   3sg deer-butcher-INTR-PAST  
   'He was deer butchering'

c.  *aapo bwe'uu-k maaso-peu-te-n*  
   3sg big-ACC deer-butcher-INT-PAST  
   ['He was [big deer]-butchering'] or ['He was deer-butchering a big one']

(1)a shows a non-incorporated transitive verb construction, where the nominal head of the object DP receives accusative case, marked with the "non-nominative" suffix *-ta*. In addition, the verb is inflected with a transitivity marker, homophous with but
etymologically unrelated to -ta. In (1)b we see noun incorporation, where the nominal root is formally attached to the verb root, with the loss of the case marker. In addition, the verb is inflected with the intransitivity marker -te. Crucially, as the ungrammatical example in (1)c shows, such constructions cannot be externally modified, i.e. these constructions do not exhibit modifier stranding. In our discussion of the other NI types it will become clear that this is what distinguishes the N-V compounding variety from the other forms of NI, which are more clearly “syntactic” in nature.

The other type of N-V compounding involves the incorporation of instrument or manner nominals, such as the instrumental prefixes of Numic, as in the Comanche examples in (2):

(2) Comanche Instrumental Prefixes in Instrumental Function

\[piH-\] ‘with the buttocks, rear’

\[a. \] [pigwain]  
\[piH-wai-n\]  
\[piH-poke-CMPL:ASP\]  
‘feeling around in the dark (for seat)’

\[b. \][Larry?a nii návukuwáa?a pihtsakíkatí]  
\[Larry-?a= nii nápukuwá-?a piH-tsaka-h/H/ka=ti=\]  
Larry-POSS I car-OBJ piH-pull, lead-TEMP:ASP  
‘I’m towing Larry’s car’

Although these instrumental prefixes usually indicate the manner or instrument with which the action of a given verb is performed, there are also cases where the prefix indicates the direct object patient of the action, as in (3):

\[1\] It is possible that there may be some ambiguity between instrumental and theme usages of instrumental prefixes in some cases. I regard these examples as “object” uses because of their glosses, e.g. ioH-tsaa ‘to
199

(3) Comanche Instrumental Prefixes in Object Function (Charney 1993: 118, 120)

a. kIH- ‘with the teeth, chin, or mouth’
   [kIgwII]  
kIH-wII
   kIH-raise
   ‘Lift your chin up!’

b. toH- ‘with the hand’, violent or completed action
   [tohtsaa]  
toH-tsaa
   toH-hold
   ‘to draw back the arm to hit’

In Chapter 8 I will claim that the crucial semantic difference between these two types of Compound NI is their underlying syntactic structure: incorporated nominals with a theme theta role derive from an underlying direct object position (i.e. sister to V), whereas nominal roots with instrumental function are derived via direct adjunction (Merge) with V. The latter process is reminiscent of the extremely productive process of deriving instrumental denominal verbs in English (e.g. hammer, nail, knife, etc.) (cf. Harley 2003).

The crucial difference between my notion of N-V compounding and the notion of NI proposed by Sapir and Mithun is that I see no formal reason to exclude instantiations of denominal verb-forming morphology from N-V compounding, since I view denominalizing morphemes as the morphological exponents (and exponence) of a verbalizing head in syntax (i.e. “little v”—c.f. Kratzer 1996, Hale and Keyser 1993, 2002, and much other work). The one difference is that denominalizing morphemes are phonologically “defective”, in the terms of Hale and Keyser (2002), and require incorporation for morphophonological reasons. The result of this lexical property is that the verbalizing head is either phonologically null or is an affix.

draw back the arm to hit’ meaning that the arm is what is held back, presumably by internal force and not by holding the held-back arm with the other arm.
Common examples of this type of denominal NI can be found in Hopi and Yaqui, but because they allow for external modifiers (doubling) they fall into the class of what I am calling “syntactic noun incorporation”. As far as I know, no Uto-Aztecan language exclusively has N-V compounding with their denominal verbs; these constructions in such a language would allow for no external modification or stranding. These would have to be analyzed as simple N-V compounding, where the relevant nominal roots are directly Merged (i.e. compounded) with the verbalizing host (i.e. the v syntactic head).

7.2.2. Syntactic NI

Unlike N-V compounding, syntactic NI allows for external modification, or “stranding”. As Rosen has noted, most languages with such a process allow for “null head modifiers” in non-NI contexts as well. Nevertheless, the link between the incorporated nominal and the null head of the phrase containing any stranded modifiers has led many theoreticians to posit a syntactic link between the two phrases (e.g. Sadock 1980, 1986; Baker 1988; Hale and Keyser 1993, 2002; etc.), thus I use the term syntactic NI for these constructions. As mentioned above, this terminology should not be taken to imply that N-V compounding (or compound NI) does not take place in the syntax; rather, it indicates only that syntactic NI leaves obvious syntactic residue, which could be captured in a variety of differing frameworks, none of which are relevant to our comparative purposes in this chapter. My syntactic analysis of compound NI will be discussed in Chapter 8; see Sadock (1985, 1991) for an alternative theoretical orientation to the relevant issues.
Examples of syntactic noun incorporation come from Hopi, where we see this process clearly with both straight-up NI (4a and 4b), with a nominal root compounding with a verbal root, as well as with denominal verbs (5a and 5b), where the verbal element is obligatorily an affix:

(4) Noun incorporation in Hopi

a. *Naat itam pu-t qatsi-yese.* (K. Hill 2003: 234 [98])
   `still we that-acc life-sit(PL)
   ‘We are still living that life.’`

b. *Nu’ pay i-t tōövu-t aw qōtō-tpe.* (K. Hill 2003: 234 [99])
   `I well this-ACC embers-ACC to.it head-roast`
   ‘I roasted this head over the embers.’

(5) Denominal verbs in Hopi

a. *Hak i-t kis-ta?* (K. Hill 2003: 234 [96])
   `who this-ACC shade-CAUS
   ‘Who built this shade?’`

   `you not what-ACC arrow-POSS-INGR
   ‘Didn’t you bring any arrows?’`

(For further examples see Hill 2003 and section 6.3.2 above).

Similar examples, with denominal verbs, are also exhibited in Yaqui:

(6) *Peo ‘iliiki-m/huevena-m kava-’ek* (Jelinek 2003: 206 [19])
   `Pete little-PL/a lot-PL horse-PERF
   ‘Pete owns a few/many horses’`

(7) *‘aapo hiva tu’ii-k te-tekil-e* (Jelinek 2003: 207 [24])
   `3sg always good-ACC RED-job-ITER-IMP
   ‘He always has a good job’`

It is crucial, though, that these kinds of stranded modifiers do *not* occur in Yaqui N-V compounding constructions (cf. example 1c above). On the other hand, the N-V
compounding constructions are of limited productivity in Yaqui—see further discussion of the analysis of these constructions in 8.4.1.

In the terms of Baker (2001), for the Hopi cases with a true verbal root in the NI constructions, this kind of incorporation can be described as "optional object polysynthesis"—the direct object of a verb may be incorporated or not, an option made available by the grammar which can be manipulated for the discourse purposes of the speaker. For the cases of denominal verb morphology, however, the incorporation is obligatory. This distinction has been the main motivation behind separating the two processes (cf. Mithun 1986, 1999; Gerdts 1998), and any account that attempts to link them will have to formalize this difference in some principled way—see discussion in Chapter 8.

In surveying the types of NI across the Uto-Aztecan family the most crucial issue is the elaboration of whether or not each language allows N-V compounding, syntactic NI, or, like Yaqui, both. From my survey it seems to be the case that all of the Uto-Aztecan languages exhibit at least one or the other in some form, although some languages may have more productive NI processes than others. However, it is not always clear from the existing grammatical descriptions of many of the languages whether or not the N-V constructions presented also take external modification, either of the stranded modifier or hyponomous argument varieties. As Rosen has suggested, however, it is clear that at least two types of NI must be described, with the possibility of external modification and/or stranding being the key diagnostic of the two types. Additionally, however, we must make clear that these constructions can co-exist in a language, and that even if a
language allows for null head modifiers, as does Yaqui, not all NI constructions may take them (cf. example 1c above).

I will have to leave for future research the status of these two types of NI with respect to the entire Uto-Aztecan family as a whole; however, see Chapter 8 for a general syntactic analysis for both types. It is hoped that the categories proposed in this chapter will be of use for Uto-Aztecanists to describe the NI processes allowed by each language. Exemplary models for work on this topic for individual languages include Hill (2003)'s discussion of NI and denominal verb formation in Hopi, based on the corpus collected in the Hopi Dictionary Project (1998), as well as Merlan (1976)'s discussion of NI and polysynthesis in Nahuatl.

In addition to the crucial question of the presence or absence of N-V compounding and syntactic NI, it is also hoped that Uto-Aztecanists will address the issue of whether null head modifiers exist in each language, and whether they can appear with noun incorporation or denominal verb formation in the form of stranded modifiers, or whether they cannot.

7.2.3. "Object polysynthesis"

The third type of NI observed in Uto-Aztecan is what I call, for lack of a better term, "object polysynthesis". As noted briefly above, syntactic NI can be viewed as a syntactic construction that Baker (2001) labels "optional object polysynthesis"—the direct object may, but need not be, incorporated into the verb. The type of NI that I label "object polysynthesis" takes syntactic NI one step further and requires object noun incorporation.
This requirement may be met either via incorporation of a nominal root into the verb, as we saw with NI in Hopi, or by an agreement marker on the verb which obligatorily agrees with the object. I follow Jelinek (1984) in regarding such agreement markers as “pronominal arguments”—these elements do not merely agree with external arguments, they themselves instantiate the arguments.

The most obvious examples of obligatory object incorporation, or object polysynthesis, come from the varieties of Nahuatl, as exemplified in (8):

(8) Obligatory Object Agreement in Nahuatl (Merlan 1976: 185)

a. askeman ti-\(^{-}\)-kwa nakatl
   never you-it-eat meat
   ‘You never eat meat’

b. na’ ipanima ni-naka-kwa
   I always I-meat-eat
   ‘I eat meat all the time’

The issue of polysynthesis is an old one in the study of the indigenous languages of the Americas, and its current usage in generative grammar generally refers to languages that obligatorily mark both the subject and the object of transitive verbs (Baker 1995, 2001).\(^2\) This is the case for Nahuatl. However, we will only concern ourselves here with the obligatory marking of the object. I follow the received view of Uto-Aztecan historical syntax, that of Langacker (1977a), in assuming that the obligatory subject agreement

\(^2\) This usage contrasts with an older notion of “polysynthesis” as referred to in morphological typology, wherein polysynthesis is applied to languages which are agglutinative to the extreme, i.e. “polysynthesis” has often been used as a “merely quantitative” notion (Sapir 1921: 136). My usage here follows Baker (1996, 2001) in employing this term in a more narrow, syntactic sense.
markers are historically derived from subject clitics, which are common throughout the family but whose manifestations vary from language to language.

Although polysynthesis has received a good deal of attention in the literature (see especially Baker 1996 and Mattissen 2003 for recent overviews), the issue of the historical origin of polysynthesis requires a good deal more investigative work in individual language families. Since full-blown polysynthesis is typical only of one sub-branch of Uto-Aztecan (Corachol-Aztecan), this language family seems to be an ideal testing ground for theories of the diachronic development (grammaticalization) of polysynthesis. These issues are addressed in detail in Chapter 9.

7.2.4. Classificatory NI

My category of classificatory NI, as a heuristic for descriptive work in comparative Uto-Aztecan, is nearly synonymous with Mithun’s Type 4 NI, but differs somewhat from Rosen’s notion of Classifier NI. The crucial factor that distinguishes classificatory NI from the previous categories is that verbs in such a construction take a hyponomous argument—i.e. an overt nominal argument that denotes some subset of the domain implied by the incorporated nominal root. This hyponymy relation is relevant both to incorporated classifiers, as in Mithun’s discussion, but also to incorporated nominal roots which are not necessarily formally classifiers in a given language. However, a hyponymy relation results from the classifying function instantiated between an incorporated nominal and its hyponomous object. In the next chapter I will present a
theory of noun incorporation that derives this relation via movement, late-insertion of morphophonological material, and co-indexation.

This heuristic category differs somewhat from Rosen’s notion of Classifier NI in that for me classificatory NI in Uto-Aztecan requires an overt object nominal that is distinct from but hyponomous with respect to the incorporated nominal, whereas the overt status of an object nominal is not relevant for Rosen. If stranded modifiers are allowed and there is no such overt noun in a given construction, then in my typology the process should be considered an instantiation of syntactic NI.

I will argue in Chapter 8, from the Pervasive Syntax Perspective (PSP), that this distinction is only illusory, in that the underlying syntactic process deriving this surface distinction is the same: syntactic head movement in each case. I present this distinction here in my effort to present a typology of NI for use in the description of these morphosyntactic processes in Uto-Aztecan because it is conceivable that some language may have only one or the other, while other languages may have both. Thus, there may be an empirical motivation to keep the two separate, descriptively at least. However, I will ultimately argue, from theoretical considerations, that classificatory NI is an instantiation syntactic NI.3

3 Hyponomous arguments must be distinguished here from cognate arguments. Cognate arguments show root identity between the incorporated nominal and what surfaces as the head of the object DP. Such an example comes from Yaqui, where a cognate argument is allowed in the context of a relative clause:

(i.) 'aapo ['uka kava'i-ta 'em'-etbwa-ka-`u-ta] kava-`ek (Jelinek 2003: 205 [14])

3sg [DET;ACC horse-ACC 2sg.P0SS steal-PERF-REL-ACC] horse-PERF

'The horse that you stole is his' ( = lit. ‘He horse-has that horse that you stole’).

Assuming the Copy theory of movement (Chomsky 1995), in most cases the lower copy is simply deleted, e.g. for pragmatic/Gricean reasons. However, there are denominal and noun incorporation constructions cross-linguistically that do require the spell-out of both copies, e.g. Hindi khaana khaa- ‘food-eating’ and gaanaa gaa- ‘song-singing’ (Klaiman 1990).
Some examples of classificatory NI in Hopi NI and denominal verb constructions were given in Chapter 6, but are repeated here:

(9) \textit{Nu’ y\textipa{n}pala-t kuy-t\textipa{ng}ta.} (K. Hill 2003: 237 [117])
\begin{itemize}
  \item I rainwater-ACC contained.liquid-put.into.container(s)
  \item ‘I put the rainwater into some containers.’
\end{itemize}

(10) \textit{Pam tsiili-t nakwa-’y-ta-ngwu.} (K. Hill 2003: 236 [112])
\begin{itemize}
  \item he chile-ACC feather.worn.on.head-POSS-DUR-HAB
  \item ‘He (the Hehey’a kachina) wears chili pepper (“chile as a feather”) on his head.’
\end{itemize}

Unlike Hopi, however, some languages may only have this function served in the context of denominal verb morphology, as in the Yaqui examples below, in (13) and (14):

(11)a. \textit{inepo chu’u-k.} b. \textit{inepo chu’u-ta hippue} (Maria Amarillas, p.c.)
\begin{itemize}
  \item 1sg dog-PERF
  \item ‘I have a dog’
  \item 1sg dog-ACC have
  \item ‘I have a dog’
\end{itemize}

(12)a. \textit{Inepo vukek} b. ?\textit{Inepo vuki-ta hippue.} (Maria Amarillas, p.c.)
\begin{itemize}
  \item 1sg pet-PERF
  \item ‘I have a pet’
  \item 1sg pet-ACC have
  \item ‘I have a pet’
\end{itemize}

(13) \textit{uka ili chu’u-ta ne=vuk-ek} (Maria Amarillas, p.c.)
\begin{itemize}
  \item det.ACC little dog-ACC 1.sg=pet-PERF
  \item ‘That little dog is my pet’ (lit. ‘I pet-have that little dog.’)
\end{itemize}

(14) \textit{inepo aa vuk-ek.} (Maria Amarillas, p.c.)
\begin{itemize}
  \item 1sg 3sg.ACC pet-PERF
  \item ‘I have it as a pet’
\end{itemize}

The issue of this classificatory NI only occurring with denominal verb formation is an issue that I will elaborate upon in the next section.

7.3. Noun incorporation in Uto-Aztecan: Results of the survey

My survey of the patterns of noun incorporation in Uto-Aztecan yielded the following results. First, the distinction between noun incorporation and derivational morphology is
typically blurred in these languages (cf. Sapir 1911, Mithun 1986, 1999) (7.3.1.). However, the full picture of noun incorporation in Uto-Aztecan will require a more standardized metric for evaluating NI types in these languages (7.3.2.). Third, the distribution of NI types in Uto-Aztecan refutes Mithun (1984)’s implicational hierarchy of NI types (7.3.3).

7.3.1. Getting over the distinction between derivational morphology and NI

Uto-Aztecan languages typically cloud the distinction between “noun incorporation” and “derivational (i.e. denominal) morphology”. This is most clearly seen in K. Hill (2003)’s comparison of noun incorporating and denominal verbs in Hopi, discussed in Chapter 6. The general argument can be summarized as follows. Since noun incorporating and denominal verbs have the same formal distribution in Hopi, each allowing stranded modifiers (adjectives, determiners and the like) as well as hyponomous arguments (i.e. displaying a classificatory function), these two morphological processes can be assumed to result from identical syntactic processes. Thus, I follow Sadock (1980, 1986), Hale and Keyser (1993, 2003), and others who link the two syntactic processes, in contrast to a tradition of separating them (cf. Sapir 1911; Mithun 1986, 1999; Gerdts 1998).

One thing that is clear about this discussion is that there is no neutral position on this issue: any stance that one takes can only be made given certain theoretical predispositions. Given my adoption of the PSP in Chapter 3, I prefer the syntactic derivation of these constructions until there is empirical reason to abandon it. However, the position that I have adopted here, based on the facts available from Hopi and other
Uto-Aztecan languages, is consistent with more functionally-oriented literature as well.

Not regarding the distinction between these two processes to be rigid and absolute accords with the grammaticalization hypothesis: affixes are historically derived from full roots, which over time become phonologically eroded, gradually evolving from free lexical items to bound elements. By hypothesis, the derivational affixes that we commonly see in Uto-Aztecan began as full verb roots that allowed for standard noun incorporation, and with a high frequency of use these verbal elements eventually eroded into affixes that actually require incorporation. Thus, the approach advocated here does not require the "noun incorporation one day and derivational morphology the next"-stance assumed in the non-gradient view held by some Lexicalists.

This theoretical stance has further implications. For example, if my account of hyponomous argument incorporation is correct, then Hale and Keyser (1993, 2002)'s NI approach to denominal verbs in English, where English unergative verbs like dance and sleep are derived by movement from an underlying object position, may be salvageable. Hale and Keyser's hyponomous argument problem (e.g. sentences such as John danced a jig), which show non-root-identity between the denominal verb and its direct object, might be justifiably considered an instantiation of "classificatory noun incorporation" (see 8.3).

---

4 There does not appear to be any Uto-Aztecan-internal evidence that the denominalizing, light verbs mentioned here have grammaticalized from free roots since the break-up of PUA. In Haugen (in prep) I take this as evidence that these morphemes were denominalizers in PUA as well.
7.3.2. A Typology for Uto-Aztecan Noun Incorporation

Because of the lack of overt discussion of the concomitants of noun incorporation in the Uto-Aztecanist literature, I propose here a typology for describing NI in these languages. These have been defined above but are summarized below in Table 7.1.

<table>
<thead>
<tr>
<th>NI Type</th>
<th>Syntactic Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-V compounding</td>
<td>Compound of Nominal root + Verb (or verbalizing affix); Valence-changing (i.e. intransitivizing); No external modification is allowed.</td>
</tr>
<tr>
<td>Syntactic NI</td>
<td>Nominal root + Verb or verbal affix; Valence is not changed; Stranded modifiers are allowed.</td>
</tr>
<tr>
<td>Classificatory NI</td>
<td>Hyponomous arguments are allowed.</td>
</tr>
<tr>
<td>Object Polysynthesis</td>
<td>Verbs require (at least) marking for direct object— instantiated either as a nominal root or a pronominal indicating the direct object argument of the verb.</td>
</tr>
</tbody>
</table>

Table 7.1: Uto-Aztecan NI Types and their characteristics

In addition to a clarification of each of these types in any given language, further discussion is needed regarding whether or not each language allows for null-head modifiers independently of noun incorporation and denominal verb formation.

7.3.3. Recasting Mithun (1984)'s Implicational hierarchy of NI types

The evidence from NI in the Uto-Aztecan languages argues strongly for a reconsideration of Mithun (1984)'s implicational hierarchy of NI types (see Chapter 6 for a detailed discussion). At the very least, the Uto-Aztecan languages suggest that classificatory NI precedes obligatory object polysynthesis (i.e. Type 4 NI >> Type 3 NI). In fact, in
Chapter 9 I will argue that "object polysynthesis" can be viewed as an extreme case of (obligatory) classificatory NI.

Secondly, however, it is unclear to what extent NI of Types 1 and 3 are required for Type 4 (classificatory) NI. It remains an empirical question, thus far unresolved, whether or not UA languages exhibit Mithun's Type 2 NI, where oblique arguments are promoted to object case position upon the incorporation of the direct object N.\(^5\) Redefining the NI types further, e.g. making the distinction between N-V compounding and syntactic NI, as I have done above, does not help the situation much, since it is not clear whether the former is required in all languages that have the latter. Yaqui, for example, has both, but some languages may only have the syntactic kind (i.e. Classifier NI without Compound NI, in Rosen's terminology).

7.3.4. Further issues in Uto-Aztecan NI

There are two further aspects relevant to NI across Uto-Aztecan. These are the presence of instrumental prefixes (7.3.4.1) and subject incorporation (7.3.4.2) in various of the languages.

7.3.4.1. Instrumental prefixes

Among the more well-known morphological aspects of the Uto-Aztecan languages are the presence of instrumental prefixes.\(^6\) Although related to NI, this phenomenon has received relatively little attention thus far in this dissertation. This is primarily because

\(^5\) DeReuse (1994) also finds Mithun's categories to be problematic for the kinds of noun incorporation displayed in Lakhota.
these affixes have a relatively limited distribution within the family, being prominent mostly in Numic and Tepiman (Langacker 1977a).

Unlike NI, where the incorporated object is typically the direct object (theme) of the verb, instrumental prefixes usually have some other thematic relation—e.g. instrument or manner. Some examples from Comanche were given above in section 7.2.1. As mentioned there, I regard instrumental prefixes as one instantiation of N-V compounding.

As is often discussed, these instrumental prefixes often have etymologically-related free nominal or verbal forms, as in the following collection from Northern Paiute (15), although not all of the prefixes have such related free forms at the synchronic stage of any given language (cf. 15b):

(15) Instrumental Prefixes in Northern Paiute (Snapp et al. 1982: 64)

<table>
<thead>
<tr>
<th>a. Instrumental Prefix</th>
<th>Related free forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. ma-</td>
<td>‘hand’</td>
</tr>
<tr>
<td>ii. ni-</td>
<td>‘talking’</td>
</tr>
<tr>
<td>iii. pa-</td>
<td>‘water’</td>
</tr>
<tr>
<td>iv. ci-</td>
<td>‘nail/pointed instr.’</td>
</tr>
<tr>
<td>v. co-</td>
<td>‘scalp/horn’</td>
</tr>
<tr>
<td>vi. mo-</td>
<td>‘face’</td>
</tr>
<tr>
<td>vii. ki-</td>
<td>‘teeth’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b. Instrumental prefixes without corresponding free roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. su-</td>
</tr>
<tr>
<td>ii. za-</td>
</tr>
<tr>
<td>iii. ta-</td>
</tr>
<tr>
<td>iv. wl-</td>
</tr>
<tr>
<td>v. to-</td>
</tr>
<tr>
<td>vi. ku-</td>
</tr>
<tr>
<td>vii. pi-</td>
</tr>
</tbody>
</table>

6 "Instrumental affixes" are also attested in other languages and language families in North America, e.g. Siouan-Catawban, Yuman, Chumash, Pomoan, Sahaptian, and others (Mithun 1999: 118-119).
In terms of diachronic analysis, however, even the prefixes in (15)b have cognate morphemes with related meanings in other Uto-Aztecan languages.

Langacker (1977a) reconstructs the following instrumental prefixes for PUA:

(16) **Instrumental Prefixes in PUA** (Langacker 1977a: 134)

*ki- ‘teeth’
*ca- ‘hand’
*ma- ‘hand’
*pa- ‘water’

As is generally assumed, I take it that instrumental prefixes such as these have their historical origin in free nouns (or verbs), which over time have become affixes through phonological and semantic erosion (i.e. grammaticalization).

That their semantic function differs from object incorporation, however, ought to indicate a distinct derivation in their synchronic syntax. Baker’s earlier theory of incorporation (Baker 1988) precluded the incorporation of nominals from adjunct positions, since such incorporation would violate the Head Movement Constraint (Travis 1984); such incorporation would involve a head moving into another head that does not properly govern it. However, in his discussion of benefactive phrases vis-à-vis instrumental phrases, Baler points out that “in some languages with NI either the instrument or the theme can be incorporated into a verb which has both” (1988: 300), as in the following example from Nahuatl (data originally from Merlan 1976):

(17)a. *Neʔ ə-panci-teteʔki ika koč*  
he 3sS-bread-cut with knife  
‘He cut the bread with a knife’  
(Baker 1988: 300 [169a])

b. *Yaʔ ki-kočillo-teteʔki panci*  
he 3sS[sic]-knife-cut bread  
‘He cut the bread with a knife’  
(Baker 1988: 300 [169b])
However, there seems to be a mistake in Baker’s gloss for (17)b, which should read: 3sO-knife-cut, since ki- is the 3rd person singular object marker. In a sense, this is a kind of “anti-Type 2 NI” construction. In Mithun’s formulation of Type 2 NI, the incorporation of a direct object nominal root allows for the “promotion” of an oblique phrase, such as a locational prepositional phrase, to become the direct object, thus allowing for the omission of the preposition (see examples in 6.2.1). In the Nahuatl case, however, it appears that an instrumental element may incorporate into a verbal construction with a pronominal direct object marker. (Merlan 1976 suggests that it is not impossible for a noun and certain modifiers to both be incorporated in a given sentence. However, for this to occur both elements would have to “constitute thematic elements in the sentence”, given that a major discourse function of noun incorporation in Nahuatl is to maintain topicality of discourse referents, free nominals being used more often for focus, and “the extent to which this normally occurs in discourse remains unclear”; Merlan was led to believe that “it may be of fairly low frequency” (1976: 188)).

In later work, though, Baker (1996) does note that certain elements, “base-generated conjoined modifiers” that have “no obvious role in the syntax”, may be attached to the verb, usually as a prefix (p. 34). According to Baker, such elements are common in Wichita and the Gunwinjguan languages, and to a limited extent Mohawk, and “are, broadly speaking, ‘modifying’, ‘adverbial’, or ‘quantificational’ in nature” (1996: 34). The prefixal nature of the instrumental affixes in Uto-Aztecan fits in with Baker’s account of adverbial modifiers in other languages, and I propose that these prefixes in
Numic (and Tepiman) are the remnants of a classificatory-style NI with full noun roots which have, through the course of grammaticalization, been reanalyzed as adverbial elements indicating the manner or the instrument with which the action denoted by the verb is performed. The synchronic analysis of these elements is that they are roots adjoined directly to the verb itself, rather than incorporating there from some other position, as do direct object nominals with the thematic role of patient (see Chapter 8.6.1.2).

7.3.4.2. Subject incorporation

A second, more serious issue for a Baker (1988)-style movement-based theory of NI is that it precludes the incorporation of true subjects. Although subject agreement affixes appear obligatorily in polysynthetic languages, such affixes are not thought to originate in an external argument position like object affixes, in theory, do. Rather, they are usually thought to be agreement elements rather than “pronouns” per se.

It should be emphasized that “subject incorporation” is only a problem for agentive subjects of transitive verbs or agentive subjects of intransitive (unergative) verbs, since the subjects of unaccusative verbs presumably originate, underlyingly, in object position. Thus, sentences such as (18)a have the underlying syntactic structure of (18)b.

---

7 The correct gloss is actually given in Merlan’s original: “3SG-it-knife-cut” (1976: 185).
8 There may also be a distinction between true object agreement and incorporated object pronominals; see discussion of this issue in 8.5.
Unaccusative Subject Incorporation in Tetelcingo Nahuatl

a. toonal-kisa
   sun-emerge
   'the sun comes out'

b. .... b'. ....
   VP
   V NP
   kisa | toonal-kisa
   emerge N
   toonal
   sun

Agentive subjects, on the other hand, with their subject projected by the higher specifier projection of vP, would require rightward and downward movement to attach to the verb, thus violating the Head Movement Constraint (Travis 1984). This appears to be exactly what we see in some examples from Nahuatl and Hopi, however.

For example, Hale and Keyser (2002) present the following examples from Hopi:

(19)a. Itam pu-t taavo-t wari-k-na
   1pl that-ACC cottontail-ACC run-K-NA
   'We flushed that cottontail rabbit out (of hiding)'
   (i.e. 'We caused that cottontail rabbit to run')

b. Itam tap-wari-k-na
   1pl cottontail-run-K-NA
   'We flushed that cottontail rabbit out (of hiding)'
   (i.e. 'We caused that cottontail rabbit to run')

In the example of (19)b, the complex predicate including the causative suffix -na is the reason for the allowing of the incorporation of tap- 'rabbit', which is, after all, the "inner subject" in this construction (i.e. tap- is the object of -na). There are other examples without this causative morphology, however.
K. Hill (2003) gives the following examples of Hopi names which were included in a 1900 census of Oraibi:

(20)a. **Hon-wari**
     bear-run(SG)
     ‘The bear ran.’

b. **Poli-wayma**
     butterfly-walk(SG)
     ‘The butterfly walked along.’

c. **Uy-hongva**
     plant-stand.up(PL)
     ‘The corn plants stood up.’

d. **Posiw-yes-va**
     magpie-sit(PL)-INGR
     ‘The magpies alighted.’

Interestingly, this data contradicts Whitely 1998’s ethnographically-informed study of Hopi naming practices. According to Whitely, Hopi names, associated with the clan affiliation of the name-giver, typically refer to specific events or instances which are only alluded to by means of the actual morphemes included in the name. These names, according to Whitely, typically suppress the subjects of the action indicated in the name. Thus, the literal meaning of a given name is “typically oblique, and not inferable from literal translation” (p. 111). For example, *lomayayva*, a Badger clan name, literally means ‘beautiful climbed (pl.)’ or ‘beautifully ascended’ (> *lolma* ‘beautiful/ly’ and *yayva* ‘climbed/ascended’), and “refers to the aesthetic splendor—in terms of color, costume, and movement—of the procession of *katsinam* at *Pasavu* coming up into Orayvi as seen from the perspective of someone standing on the mesa” (p. 111). The suppression of the subject in these cases seems to be linked to their association with *katsinam*, so it is
conceivable that the subjects in the names in (20) are more literal than reference to *katsinam* would indicate. However, the literal syntactic formulation of these names does not render the context of their utterance, revealed by the names themselves as “tiny imagist poems” (Whitely 1998: 111—see also Basso 1996), any less opaque.

Hill also points out that similar agentive subject-incorporating names occur in Classical Nahuatl, as in (21):

   ‘The eagle descended.’

   b. *Mo-teuc-zōmah* (In tēuctli ōmozōmah.)
   ‘The lord frowned in anger.’

As Hill points out, the use of such subject incorporation in the naming practices of two such widely divergent languages in the Uto-Aztecan family brings up the possibility that this is an ancient practice common to the Uto-Aztecan languages. It would be interesting to survey the other languages to ascertain how common this currently is in the family, or how frequent such names are in colonial documents. A full survey, of course, goes beyond the scope of the present work, although I have not come across other examples in my review of the extant grammatical descriptions.

Most of the subject-incorporating structures that have been discussed thus far involve verbs of motion, which conceivably could lead to an unaccusative object analysis of these constructions. However, the example in (21)b, glossed as *The lord frowned in anger*, is not such a verb, and seems to be a more clear case of agentive action. On the other hand, the fact that these are all names leads to the question of whether or not some non-sentential analysis might be best to account for them.
Returning to the issue of the syntax of these constructions, similar phenomena have also been reported in other languages, including Koyukon Athabaskan (Axelrod 1990) and Lakhota (de Reuse 1994), and such constructions need to be taken into account in a general theory of syntactic word-formation.

The solution to this puzzle that I present in Chapter 8 is that such instances of “agent subject incorporation” cannot be derived via head movement, as Baker’s theory suggests. However, as has been repeatedly emphasized in this work, there is another way to generate an N-V “incorporation” construction, and this involves simple compounding. Thus, in these examples I hypothesize that the construction is simply the result of the Merging of a nominal root into the verbal position, rather than derivation by head movement. This compounding leads to an idiomatic interpretation wherein the incorporated nominal is taken to be the subject of its verb. As Mithun (1984) has amply illustrated with examples like the English alligator shoes, such compounding does not force a particular interpretation, so idiomatization can lead to agentive readings in some cases, and to direct object readings in others.

This should not be possible with “syntactic NI”, where the incorporating N is derived from underlying object position leaving a trace or allowing for the insertion of a hyponymous argument. Thus, ”subject incorporation” constructions should not be able to have “stranded modifiers” for the incorporated subject. This seems to be the case from the examples that I have found in the literature (see 8.6.2. for further discussion).
7.4. Noun incorporation in Uto-Aztecan: The data

In this section we will survey the noun incorporation and denominal verb constructions for each of the major Uto-Aztecan sub-groups. This survey will show the evidence for each of the NI types developed above where it exists for each language, and will point to areas where questions still remain.

This section is divided into subsections based on the sub-groupings of Uto-Aztecan given in Figure 2.1, which are arranged roughly in order from north to south.

7.4.1. Numic: Comanche

The Numic languages are sub-divided into three groups, as shown in Figure 2.1. These languages are famous for their instrumental prefixes. In addition, each of these languages show various NI and denominal verb constructions. My focus in this section will be on Comanche, a Central Numic language.

Comanche has classical compounding NI, where the N-V compound forms an intransitive verb, and where the incorporated noun is the logical object of the verb. The incorporated noun appears in its bare root form, with no inflection. Contrast the incorporated form in (22)b with the unincorporated form in (22)a:

(22)a. \[\text{kikari inni saahuve nimákihutu?i}\] (Charney 1993: 123 [36])
\[\text{kikati inni sahupa-e ni-maka-hu=(2)-tu?i}\]
EXPC you soup-obj. me-give, feed-INTN:ASP-UR:ASP
'I thought you’d give me some soup'

---

9 I follow Charney (1993)’s transcription conventions and morpheme glosses. The phonetic representation is given in brackets, while the morphophonemic representation is given below in italics.
b. [kikari inni nisåhuvanakhutu?i]  
\[\text{kikati inni ni-sahupa-maka-hu=(2)-tu?i}\]  
EXPC you me-soup-give, feed-INTN:ASP-UR:ASP  
‘I thought you’d give me some soup’

As emphasized by Mithun (1984), these kinds of verbs are often used for “habitual activities” (e.g. 23b vs. 23a), and some NV compounds have non-compositional, idiomatic meanings (24).

(23)a. [uhka nii sone marikai]  
\[\text{uHka nii sona-e ma-tika-i}\]  
‘I finished that quilt’

b. [tiasi nii sona?ai?etI]  
\[\text{tiasi= nii sona-ai-?e-t?=}\]  
‘I also do quilting’

(24)a. [pukutsaka]  
\[\text{puku-tsaka}\]  
‘to testify’

b. [pahavi]  
\[\text{paa-hapi}\]  
‘to swim’

c. [tomoyake]  
\[\text{tomo-yake}\]  
‘to thunder’

However, it is clear that not all NI constructions are habitual or idiomatic, since compositional meanings can also be derived from these constructions:
Charney does not mention and provides no obvious examples of the possibility of NI constructions leaving stranded modifiers—the examples that she provides for NI, and object pronouns, typically involve indefinite reference (cf. the non-incorporated object in 22a above).

There are no clear examples of classificatory NI in Charney (1993)’s grammar. One possible way to explore this issue further would be to elaborate upon the use of object prefixes. For example, if a verb takes a third person object prefix and an overt third person direct object DP, then this would constitute classificatory NI according to my typology of NI. Such a construction would translate as the paraphrase ‘Subject it-verbs DP’. Charney does give many examples of the object pronouns being used with a direct object incorporated noun, however, and these yield benefactive readings:

(26)a. [nihúúva?aiki?i u] (Charney 1993: 102 [3])
\[ni-hupa-?ai-ki \quad u\]
me-coffee-make-BEN-CMPL:ASP she
‘She made coffee for me’

\[tahi-ta ?o-?ai-ki=i \quad nii\]
us.DU.INCL-pound.meat-make-BEN-CMPL:ASP I
‘I made pound meat for the two of us’

---

10 There are twelve possible object pronouns, four each for singular, plural, and dual. It is not clear whether all of these can be bound to the verb as a prefix, or if they can all be free pronouns. Charney’s examples appear to be bound (pp. 101-2).
In addition to these N-V compounds, Comanche productively allows N-N, V-V, and modifier-N compounds.

Related to NI is the presence of unspecified argument prefixes, including indefinite object markers *ma-* and *ti-. According to Charney, *ma-* tends to be more definite in reference than *ti-, and is also more often used for humans, whereas *ti- is more often used with inanimates. (There is also an indefinite subject marker, *ta-=*, which “almost always occurs in subordinate clauses or as nominalizations; *ta= - seldom occurs with the main verb of the sentence”, p. 129). If verbs with the object prefix could also take an external direct object then this construction would also be equivalent to classificatory NI—however, Charney gives no such examples.

As to the use of derivational morphology for denominal verb-formation, Comanche has at least the following derivational morphemes:

(27) Comanche denominalizing suffixes (Charney 1993: 204-208)

- *pai*  
  'have'

- *ka=ti=*
  'have'

- *?ai*
  'make, do'

- *tu*
  'to marry' (lit. 'get, acquire an affinal relative')

- *piHka*
  'to become'\(^{11}\)

According to Charney, “[g]enerally, -*pai* indicates relatively permanent possession of nominal objects, while -*ka=ti= is used for objects that are less permanently possessed, as well as for the possession-attribution of qualities” (pp. 204-5). Possessive -*pai* is examplified in (28), and -*ka=ti= in (29); for -*?ai* ‘make, do’ see (23b) above:
(28)a. [suRIse nikanu tammati sonipl=paï] (Charney 1993: 205 [78])
   suti=se n-i-kinu tamma=ti soni=pi=pai
   that.one-CNTR my-a.lot-OBJ grass, hay-ABS-pai
   father’s.father
   ‘My grandfather had hay (for horses)’

b. [innha nakihtava?ivai] (Charney 1993: 205 [77])
   inni-ha nakiHtapa ?i-pai
   you-WH pecan-pai
   ‘Do you have any pecans?’

(29)a. [kehéna nii nakihtava?ikatI] (Charney 1993: 205 [76])
   ke-hena nii nakiHtapa ?i-ka=ti=
   NEG-something=OBJ I pecan-ka=ti=
   ‘I don’t have any pecans’

b. [kahnikati nii] (Charney 1993: 100 [1])
   kahni-ka=ti=nii
   house-have I
   ‘I have a house’

c. [ahpikati mikwl] (Charney 1993: 100 [1])
   aHpi-ka=ti=mikwih
   father-have you=DU
   ‘You two have a father’

The pragmatic difference between -pai and -ka=ti= is wonderfully exemplified by the following example (Charney 1993: 205 [71]):

(30)[wahati nii monahpikatl. tiasise nii wahati huutsipiapivai.] (Chamey 1993: 205 [78])
   waha=ti nii monaH=-pi-ka=ti= tiasi=se nii waha=ti huutsi=pia-pai
   2.NOM.OBJ I son.in.law-ABS-ka=ti= also-CNTR I two-NOM.OBJ daughter.in.law-ABS.pai
   ‘I have two sons-in-law and two daughters-in-law’.

11 According to Charney, this suffix means ‘to become’ ‘in an undesirable sense, as in ‘to become angry’ or ‘to become old and feeble’. There is also a verb root, naha ~ na’la which is used to express the idea of ‘becoming’.
According to Charney, in Comanche culture "sons-in-law are relatively impermanent, and the term for them is suffixed with –ka=ti=, whereas daughters-in-law, because the grandchildren remain with them, are relatively permanent 'possessions' and thus have the –pai suffix" (p. 222).

Examples of the use of –tu ‘to marry’ are given in (31):

(31)a. [inni kuhmarun] (Charney 1993: 206 [83])

*inni ku*ma-tu-n
you male-tu-CMPL:ASP
‘Did you get married?’ (female address)

b. [inni kwí?itun] (Charney 1993: 206 [84])

*inni kwí?i-tu-n*
you female-tu-CMPL:ASP
‘Did you get married?’ (male address)

An example of the use of -piHka ‘to become’ is the following:

(32) [tsukuvihkai u] (Charney 1993: 199 [31])

*tsuku-piHka-i u*
old.man-become-CMPL:ASP he
‘He’s getting up in age’ (lit. ‘He’s becoming an old man.’)

Examples of stranded modifiers appear in (30) above; it is not clear if these constructions can also take hyponomous objects in Comanche.

We also see stranded modifiers in other Numic languages, as in the following examples:

(33) **Southern Paiute** (Sapir 1930:134)

*wa’q-utcant qava ‘x:A*
two-OBJ-preterite-I horse-GET
'I received two horses'
(34) **Northern Paiute**

*usu* hici puku-ga-si iwa puku-dua

he few horse-HAVE-SUBR many horse-INCEP

'He had just a few horses; then his horses became many.'

(35) **Tümpisa Shoshone**

*Nü* sakwaapitūnna puipa'ē.

I green-OBJ eye-HAVE

'I have green eyes'

(36) **Gosiute Shoshone**

*Ni* wahatti kahnikantān

I two-OBJ house-HAVE

'I have two houses'

In addition, hyponomous arguments are also attested with denominal verbs in other Numic languages, as in the following examples:

(37) **Tümpisa Shoshone**

*Nūmmū* so'oppūh putish pungkuaimmpūhantū

we(exc) many burro pet-HAVE-HAB-PAST

'We used to have many burro pets'

(38) **Gosiute Shoshone**

*Isapaippūh* sukka ponaiha taipai

Coyote that-OBJ Mouse-OBJ brother-HAVE

'Coyote has Mouse for a younger brother'

Thus, modifier-stranding and the classificatory function are typical of denominal verbs in at least some Numic languages.

Finally, Comanche exhibits the range of instrumental prefixes that is expected in a Numic language. Some of these instrumental prefixes can act as object of the verb, as in the following examples (repeated from 3 above):
(39) kiH- ‘with the teeth, chin, or mouth’

[kiH-wii]

kiH-raise
‘Lift your chin up!’

(40) toH- ‘with the hand’, violent or completed action

[toH-tsaa]

toH-hold
‘to draw back the arm to hit’

However, instrumental usage is also attested:

(41) mu- ‘with the nose, lips, front’

[muH-wai]

mu-poke
‘to poke with the nose’

The full range of instrumental prefixes listed by Charney are given in (42); for exhibition of the use of these prefixes in sentential context, refer to Charney (1993: pp. 117-123):

(42) Comanche Instrumental Prefixes

a. kiH- ‘with the teeth, chin, mouth’

b. kuH- ‘with heat, fire’

c. ma- ‘with the hand’, also a generalized instrumental

d. mu- ~ muH- ‘with the nose, lips, front’

e. niH- ‘verbally’
f. piH- ‘with the buttocks, rear (e.g. of a car)’
g. si-= ‘with cold’
h. siH- ‘with the foot’, in a violent motion

i. su= ‘with the mind’, mental activity

j. taH- ‘with the foot’

k. toH- ‘with the hand’, violent or completed action

l. tsaiH- ‘with the hand’ (extended to hand tools)
m. tsiH- ‘with a sharp point, with the finger’
n. tsox- ‘with the head’
o. wiH- all-purpose instrumental
In sum, although stranded modifiers and hyponymous arguments occur with
denominal verbs in Numic, it is unclear from my survey whether these languages allow
for the same with noun incorporation. In his grammar of Túmpisa Shoshone, Dayley
(1989) regards the denominalizing morphemes as “regularly incorporating verbs” that
obligatorily incorporate their objects (p. 90), thus apparently sharing my view that the
verbal root/denominalizing affix distinction is a blurry one. However, according to
Dayley, noun incorporation, outside of these denominal verb constructions, is “sporadic
and unproductive” (p. 91).

7.4.2. Takic: Cupeño

Noun incorporation and denominal verb formation in Cupeño must be understood within
the overall context of person-marking in this language.

With respect to subject (agent)-marking, subject affixes, which indicate person and
number, are required on the verb in the past tense, whereas other tenses do not require
(and do not allow) such marking—see J. Hill (2003b) for discussion of the details which
go beyond the purposes of this chapter.

In reference to object marking, which is more relevant to our discussion here, object
agreement may be indicated on the “verb complex” in both past and non-past contexts.
This object “pronominal” marking precedes subject marking in the past tense. Although
not necessarily obligatory, these pronominal elements usually appear for non-third person
singular arguments. For a variety of reasons Hill regards these object markers as “pro-
clitics” rather than “pronouns” or “pronominal affixes”.
First, the object marker is ordered before the subject marker in non-past tense forms, which is relatively unusual in Uto-Aztecan (cf. Nahuatl, a polysynthetic language, where the subject prefix precedes the object prefix or incorporated noun—see Chapter 9). Secondly, these object markers never appear adjacent to the transitivizing light verb \(-in\), which is the head of vP (Barragan 2003). Additionally, these elements are never stressed, even when appearing with roots with no stressable affixes, unlike subject prefixes and tense/aspect/mood suffixes, which do accept stress. Finally, in Hill’s corpora she notes hesitations, or “disfluency”, between the object pronominal and rest of the verb construction, but not with subject or other affixes which attach to the verb.

Thus, the relationship between the verb and its pronominal direct object does not show the degree of phonological “fusion” that is otherwise expected from other languages with obligatory object pronominal prefixation on its verbs, such as Tohono O’odham. Thus, whereas the clitic status of pronominal elements might seem to indicate that NI, at least of pronouns, is “weak” in Cupeño, the opposite conclusion will be reached for Tohono O’odham in section 7.4.4 below.

However, these pronominal clitics can be coindexed with external argument DPs, thus displaying the crucial characteristic of classificatory NI, as in the following examples:

\[(43) \quad (\text{J. Hill in press: Ch. 4 [7a]})\]
\[a. \text{Mu} = \text{ku}'\text{ut} \quad \text{aye} \quad \text{pe-} \text{na}'\text{aqwa-nm-i} \quad \text{mi} = \text{kwaw-pe-n} \]
\[\text{AND} = \text{REP} \quad \text{THEN} \quad 3\text{S-CHILD-PL-O} \quad 3\text{PLO} = \text{CALL-3S-IN} \]
\[\text{“And then it is said he called his children.” (Faye Creation 119)}\]

\[\text{It should be noted, however, that object marking also precedes subject marking in Navajo and other Apachean languages in Southern Athabaskan (Jelinek 1989, Willie 1991, Rice 2000), where these elements are usually regarded as affixes. The free root-clitic-affix continuum is only tangential to the points that I will develop below.}\]
b. Pe-$huun-i pi=kulu-lu-pe-n-ngiy (J. Hill in press: Ch. 4 [7b])
   3S=HEART-O 3S=DRAG-DUP-3S-IN-MOTG
   “He went away dragging his heart.” (RN Creation 123)

These proclitics can also code for indirect and benefactive objects, which incorporate in lieu of direct objects when they co-occur in the same sentence, as in (44) and (45), respectively:

(44) Indirect object (J. Hill in press: Ch. 4 [7c])
   a. Qay=’ep hi-sh e-ach-i chimi=’uni-qqa
      NOT=2SERG WHAT-NPN 2S-PET-O 1SO-SHOW-IN-PRS
      ‘You did not show us your pet.’ (Faye KP 139 187)

(45) Benefactive object (J. Hill in press: Ch. 4 [7d])
   b. Em-em=qwe=me chimi=mixaan me chimi=meqen-max hunwe-t pe’
      YOU-PL=CAN=2PLERG 1PLO=DO.AAN AND 1PLO-KILLS-BEN BEAR-NPN DET
      aye chimi=tul-qqa
      NOW 1PLO OB-FINISH-PRS
      ‘You (pi) must do something for us, and kill for our sake the bear who is now finishing us off.’ (Faye KP 151 217)

For these latter cases, where the agreement of the proclitic is triggered by arguments that are not complements to V, I would argue that the head of the direct object phrase does not incorporate, but rather remains with the DP.

With respect to other types of NI, Hill reports (personal communication) that Cupeno does not have N-V compounding (classical noun incorporation), although one can identify etymologically the remnants of such a process in such fossilized morphology as instrumental prefixes, which are also not productive in the synchronic language.

As expected from the larger discussion, Cupeno has suffixes that derive denominal verbs. Like non-derived verbs in Cupeno, denominal verbs fall into classes based on the verbalizing stems they take: -in, -yax, or φ.
Some examples of the inchoative use of denominalizing suffix -chu are given below:

(46) Inchoatives in -chu (J. Hill in press: Ch. 7 [41])

- a. pi’myki-sh ‘ghost’ → pi’myk-chu ‘to turn into a ghost’
- b. ki-sh ‘house’ → ki-chu ‘to dwell, stay in a place’
- c. naxani-sh ‘man’ → naxan-chu ‘to grow old, of a male’

Other suffixes deriving nouns are -tu, -lu ~ -lyu, or -chu. Hill’s analysis of these denominal verb morphemes in Cupeno is that they are underlyingly -tu, with surface allomorphs -lu ~ -lyu and -chu. These allomorphs may be the remnants of “final features” traced back to PNUA, but which are not longer active in Cupeno. However, the same consonantal ending appears with Non-Possessed Noun (NPN) suffixes (i.e. the Uto-Aztecan “absolutive”) as well, so Cupeno nouns can be divided into classes based on which consonantal ending appears. For our purposes, we will regard the suffix as -tu and not concern ourselves with the conditioned allomorphy that alters the surface form.

Hill states that “[t]hese derivations yield verbs with a variety of meanings that can be loosely translated as ‘having, having the quality of’,” and in interlinear glosses she neutrally refers to the morpheme as “VB” (verbalizer). I take it that this morpheme is the exponent of a light verb (“v”) head into which nominal roots incorporate to form denominal verbs, as I will discuss in Chapter 8. As Hill points out, this light verb is probably a reflex of the common Uto-Aztecan suffix -tV, which appears, in different functions, throughout the family.

Here I will separate the noun + -tu complexes given by Hill into groups based on their apparent semantic uses, as ascertained through their glosses; examples that may fit equally well into more than one category are given in both with question marks to
indicate their marginal status. (All of the following examples come from J. Hill in press, figure 45: I retain her original lettering for each example).

(47) -tu ‘have, possess’
   a. ash-lyu ‘have a dog’ (> -ash ‘pet’, achi-ly ‘cow’)
   d. hak-lu ‘to starve, experience famine’ (> hakwiga ‘be hungry’)
   j. mukwi-lyu ‘have sores’ (> -myk’i-ly ‘sore’)

   (48) -tu ‘have the qualities of’
   c. ew-lu ‘to be initiated at puberty, for girls’ (> ewe-l ‘blood’)
   f. i’is-lyu ‘tell lies’ (> isi-ly ‘coyote’, i’isi-ly ‘liar’)
   i. meme-lu ‘speak English’ (> meme-m ‘Whites’)
   k. mu-lu ‘to lead, go first’ (> -mu-s ‘nose’)
   l. pa-lu ‘to be watery’ (> pa-l ‘water’)
   n. pew-lyu ‘to be a friend, be friendly’ (> -pew ‘friend’)
   o. push-lyu ‘to look like someone’ (> push-s ‘eye, face’)

   (49) -tu ‘use’
   b. ay-lyu ‘shake a rattle’ (> ayi-ly ‘tortoise-shell rattle’)
   e. ika-l (?) ‘to knit’ (> ika-t ‘carrying net’)
   k. mu-lu ‘to lead, go first’ (> -mu-s ‘nose’)

   (50) -tu ‘marry’
   g. kuung-lu ‘get married, speaking of a woman’ (> -kuung ‘husband’)

   (51) -tu ‘become’
   r. ye-lu ‘become a mother’ (> -ye-s ‘mother’)

   (52) -tu ‘be’
   h. liimpyu-lu ‘be clean’ (> liimpyu ‘clean’)

   All of the above forms can be straightforwardly accounted for if we posit incorporation of the nominal into a light verb head v, which contains the semantic meaning attributed to the N-v complex (e.g. HAVE, USE, etc.).

   Some forms of this type can lead to further derivation with the causative suffix -nin, which leads to a causative reading (J. Hill in press: Ch. 7 [46]):
(53)  -tu + -nin ‘make, cause’

g. qilyiq-tu’-nin ‘to hurt’ (> unattested qilyiq-tu ‘be spicy, hot’; qilyiq ‘spicy, hot’)
h. wi-tu ‘get fat’, wi-tu’-nin ‘make fat’, (> wiwat ‘fat’)
p. tash-lyu ‘gamble’ (> tash- ‘to break into small pieces, to crack acorns,’ tash-ni ‘to hatch’, tachil- ‘to split, crack’); perhaps in reference to the small gambling tokens that are spread out in front of the player)
q. tew-lyu’-nin ‘to cause to have a name’ (from unattested tew-lyu ‘to have a name’; tewe-l ‘name’)
s. yu-lyu’-nin ‘to blame’ (from unattested yu-lyu ‘to have head, responsibility? ’; yu-’el ‘head, hair’)

Hill also lists some other suffixes, which she refers to as “minor suffixes deriving verbs” because of their limited distribution (pp. 48-55), which we will not consider here.

Although most of the examples of denominal verbs in Hill’s grammar are not accompanied by sentential exemplification, Cupéñö does allow for hyponymous objects with denominal verbs, as in the following example:

(54)  tuku=’ep ne-âsh-lyu awá-l-i (Jane Hill, p.c.)
yesterday=r ls-pet-HAVE dog-NPN-OBJ
‘Yesterday I had a dog’

It is less clear whether or not modifier stranding can occur with denominal verbs; according to Hill (personal communication) if they can they are rare.

7.4.3. NUA Isolate: Hopi

The facts regarding denominal and noun-incorporating verbs in Hopi are presented in great detail by K. Hill (2003), and were discussed at length in 6.3. As was pointed out there, both denominal and noun-incorporating verbs leave stranded modifiers and show
instantiations of classificatory NI. An open question that remains is whether or not Hopi has intransitivizing N-V compounding, like the (unproductive) process we see in Yaqui.

There is an additional area of interest regarding Hopi noun incorporation that I would like to discuss in this section. There are a limited number of verbs that K. Hill calls “pronominal verbs” because they require a pronominal prefix indicating their direct object. According to Hill, this class of verbs contains only five verb roots, listed in (55):  

(55) Hopi “Pronominal Verbs” (K. Hill 2003: 221 [23])

a. (a) 'tsiva ‘behave in a manner in accord with the nature of’
b. (âa) sala ‘spread all over, through’
c. (âa)sawva ‘meet and pass in opposite directions’
d. (âa)tsavala ‘scatter’
e. (a)ptu ‘become enough for’

According to Hill, “the first and second person objects of such verbs are not expressed beyond the prefix itself, but the third person forms of this prefix may be in agreement with a separate object expression” (2003: 221). That is, third person pronominal verbs may take hyponomous arguments.

Some examples are given in (56) (from K. Hill 1998: 881):

(56)a. Tiyōya paavay àa-pungyala
   little.boy older.brother 3rd.sg.-keep.wanting.to.hang.around.with
   ‘The little boy wants to hang around with his older brother.’

b. Nu’ nuunukpàntuy amìu-piyna
   1sg evil.ones-ACC 3rd.pl-make.leave
   ‘I made the bad guys go away.’

---

13 The third person singular pronominal prefix appears in parentheses—this may be replaced by one of the other pronominal prefixes.
Although the Hopi Dictionary Project (1998) lists more than the five verb roots discussed by Hill (2003), it seems clear that this class of verbs is highly limited in this language.

The analysis of these and other hyponomous argument facts that I will sketch in Chapter 8 is that the incorporated element, here pronominal, derives from the same syntactic position that its direct object argument appears—i.e. sister to V. My analysis proposes that hyponomous arguments may fill in the "trace" (copy) left behind after movement, a consequence of the Late Insertion of lexical material into syntactic configurations.

In this limited class of verbs, there is a lexical requirement that the incorporated element be the highest possible element in the semantically-based hierarchy in which roots are arranged—i.e. a deictic pronominal element. That Hopi displays a quasi-pronominal argument subsystem within its grammar offers possible support for the gradual development of a full-blown pronominal object "parameter", wherein all transitive verbs have this requirement. We turn now to discuss one such language in Uto-Aztecan: Tohono O'odham.

7.4.4. Tepiman: Tohono O'odham

Tohono O'odham, formerly known as Papago, is a pronominal argument language (Jelinek 1984, 2001) that obligatorily requires object prefixes on all transitive verbs. This requirement is similar to that found in polysynthetic languages, where verbs must be marked for both subject and object (Baker 1996, 2001). However, Tohono O'odham is not, strictly-speaking, polysynthetic, since subject agreement marking in O'odham is
obligatory on the second position auxiliary element (Zepeda 1983, Jelinek 1984), rather than on the verb itself, and syntactic noun incorporation does not occur on most verbs.

I follow Jelinek (1984, 2001)'s proposal that these affixal "agreement" elements in this language constitute the actual arguments of the verb—the direct object prefix attached to the verb is the verb's direct object. One consequence of Jelinek's analysis is that full DP's, which may co-occur with such prefixes, are not themselves in argument positions, but are rather adjunct (A) elements that may be freely moved around the sentence for the discourse-functional purposes of the speaker. Thus, in Jelinek's framework, the obligatory argument prefixes allow for non-configurationality among free nominal phrases in O'odham sentences.

In my account of these elements in Chapter 8, however, the full DP object arguments are in object position, at least underlingly. These constructions, with an incorporated pronominal and an external hyponomous object, are an instantiation of classificatory noun incorporation. The unique thing about O'odham, with respect to the other Uto-Aztecan languages that we have been considering, is that this classificatory NI is obligatory in this language.

The object prefixes of O'odham are given in (57)a; these contrast with the optional free pronouns (57)b (from Zepeda 1983: 35):

(57)a. O'odham Object Prefixes

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person:</td>
<td>ſ-</td>
<td>t-</td>
</tr>
<tr>
<td>2nd person:</td>
<td>m-</td>
<td>em-</td>
</tr>
<tr>
<td>3rd person:</td>
<td>Ø-</td>
<td>ha-</td>
</tr>
</tbody>
</table>

(57)b. O'odham Free Pronouns

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person:</td>
<td>'a:ňi</td>
<td>'a:čim</td>
</tr>
<tr>
<td>2nd person:</td>
<td>'a:pi</td>
<td>'a:pi</td>
</tr>
<tr>
<td>3rd person:</td>
<td>hegai</td>
<td>hegam</td>
</tr>
</tbody>
</table>
The use of the free pronouns as direct objects is optional, and they can be used for focus, contrast, and other pragmatic effects.

With respect to noun incorporation, Jelinek (2001)'s proposal for her “pronominal argument parameter” suggests that pronominal arguments should preclude the possibility of NI with free verb roots, since these verbs require expression of their direct objects to be pronominal in nature. Given the syntactic representation defended in Chapter 8, these verbs require the insertion of a classifier element, which in this language involves the personal deictics. As far as I know, Jelinek’s proposal bears out—NI does not seem to occur in O’odham.

O’odham does have denominal verb structures, however, which, in the terms employed in this dissertation, require the incorporation of their nominal objects. Zepeda (1983) lists the following suffixes:

(58) Denominal Verbalizers in O’odham (Zepeda 1983)

- mad ‘use’ or ‘put’
- pig ‘remove X’

The semantics of -mad can typically be summed up under the general rubric of “use”, or “do with X that what one characteristically does with X”, in most instances:

(59) a. cu’i ‘flour’ \(\rightarrow\) cu’imad ‘adding flour to, flouring’
    b. ko’okol ‘chile’ \(\rightarrow\) ko’okolmad ‘adding chile to’
    c. ‘on ‘salt’ \(\rightarrow\) ‘onmad ‘adding salt to, salting’
    d. sitol ‘honey’ \(\rightarrow\) sitolmad ‘adding honey to’
    e. jewed ‘dirt’ \(\rightarrow\) jewedmad ‘getting dirty’
    f. siswui ‘spit’ \(\rightarrow\) siswuimad ‘spitting on something’

Some of these forms must derive from a complex underlying structure, since some can still take a direct object:
I propose that in these cases the underlying structure is similar to the location-verbs analyzed by Hale and Keyser (1993), with an abstract postpositional element allowing a reading that can be paraphrased as “CAUSE X to be in/on Y”.14

Saxton (1982) gives the following additional denominalizing suffixes:

(62) Other denominal verb morphemes in O'odham (Saxton 1982: 163-5)

| a.  | -t | MAKE |
| b.  | -mad | USE     | (apply N) |
| c.  | -cud | APPLIC | apply (inanimate) N, cause to be, or treat like (animate) N |
| d.  | -dag, dag | PUT ON |
| e.  | -g | EXIST |
| f.  | -gid | APPLIC | shake N |
| g.  | -hun | APPLIC |
| h.  | -mid/-op | MOTION | of purpose, to go for N |
| i.  | -pig | PRIV | to remove |

Similar morphemes serving similar functions are also found in Northern Tepehuan (Bascom 1982) and Southeastern Tepehuan (Willett 1991).

In addition to the denominalizing morphemes above, Saxton (1982) also lists a collection of morphemes that indicate some action modified by (usually) the instrumental use of the incorporated nominal—see Saxton (1982: 163-166) for examples.

---

14 This suffix is not fully productive—not all nouns can be used in this way. For example, some zero-derived unergative verbs “block” the addition of this suffix to a nominal root: sudagi ‘water’ \(\rightarrow\) *sudagimad (cf. wa: ‘to water’).
There are also some additional, rather idiosyncratic denominalizing morphemes (Saxton 1982: 163):

(63) **Additional denominal verb morphemes** (Saxton 1982: 163)

a. -giw  move $N$ constantly
b. s=. .-hog to be displeased by
c. -hog to be expected to
d. -šid to mimic the action
e. -šan, -ša-d to act along a surface
f. -wua to bump or touch $N$
g. -s RESULTATIVE, PASSIVE ("2nd rank argument replaces the highest")

Tepiman denominal verbs can also take hyponymous arguments:

(64) **Tohono O’odham**

\[
\begin{array}{c}
 n-t \quad wo \quad ha’i \quad kii-ki-t \\
 I-TNS \quad FUT \quad some \quad RED-house-MAKE \\
 ‘I’m going to build some houses’
\end{array}
\]

(65) **Southeastern Tepehuan** (Willett 1991: 63 [135])

\[
\begin{array}{c}
 Na \quad gu’ \quad cham \quad ja-via’ \quad gu \quad vác, \quad cham \quad mat \quad va’ \quad gu \quad quis \\
 SUB \quad but \quad NEG \quad 3p-have \quad ART \quad cows \quad NEG \quad know \quad then \quad ART \quad cheese \\
 ‘Because he doesn’t have any cows, he doesn’t know (how to make) cheese’
\end{array}
\]

To conclude this section, Tohono O’odham is a typical Tepiman language that fulfills Jelinek’s criteria for a pronominal argument language. Noun incorporation is not productive, except in those instances where a verbal affix requires incorporation of a nominal (i.e. in denominal verbs). Standard transitive verbs require pronominal marking for objects, making O’odham similar to a polysynthetic language in this respect; however, subject pronominal marking is limited to the second position auxiliary position. I will provide unified syntactic analysis of direct objects in O’odham and other Uto-Aztecan languages in Chapter 8.
7.4.5. Taracahitic: Yaqui

Yaqui has each of the types of noun incorporation that I have proposed for the Uto-Aztecan languages, and indeed, it was the study of this language that motivated the search for similar patterns elsewhere in Uto-Aztecan.

The process of N-V compounding can be clearly illustrated with examples such as the following, already discussed in Chapter 6:

(66)a. *aapo maaso-ta peu-ta-k
3sg deer-ACC butcher-TRAN-PERF
‘He butchered a deer’

b. aapo maaso-peu-te-n
3sg deer-butcher-INTR-PAST
‘He was deer butchering’

c. *aapo bwe’uu-k maaso-peu-te-n
3sg big-ACC deer-butcher-INT-PAST
[*‘He was [big deer]-butchering’] or [*‘He was deer-butchering a big one’]

The ungrammaticality of (66c) indicates that external adjectival modification of the incorporated N in these constructions is impossible. This is also true for possessive modifiers:

(67)a. Peo Huan-ta maso peu-ta
Peo Huan-POSS deer butcher-TRANS
‘Peo is butchering Huan’s deer’

b. *Peo Huan-ta maso-peu-te
Peo Huan-POSS deer-butcher-INT
‘Peo is butchering Huan’s deer’

c. Maso-peu-ti-wa-k.
deer-butcher-INT-PASS-PERF
‘Deer-butchering occurred.’
d. *Huan-ta maso-peu-ti-wa-k
   Huan-POSS deer-butcher-INT-PASS-PERF
   *[Huan's deer]-butchering occurred' (= 'Huan's deer were butchered')

As Rosen (1989) points out, most languages that allow for the stranding of modifiers do so whether or not there is noun-incorporation, and Yaqui is one of these languages. However, under the Lexicalist account of this phenomenon it is unexplained why stranding is not possible in examples such as these. What I argue in the next chapter is that two mechanisms of incorporation lead to the distinct syntactic patterns that we see with NI: N-V compounding is the result of Merge (i.e. compounding), where the V takes a "nominal" root complement, while "syntactic NI" is derived via syntactic head movement from a full DP (or NP) complement.

It should be emphasized, however, that the process of N-V compounding does not seem to be fully productive in Yaqui. It is not the case that all verbs can freely incorporate all likely object nominals. As stressed by Mithun (1984), cross-linguistically, N-V compounding of this type is often given only to situations involving "name-worthy activities", and this seems to be the case in Yaqui. My own experience with noun incorporation in Yaqui is limited to the verb root *peu*- 'butcher', and it can only be applied to largish animals which are typically butchered (e.g. deer and pigs), rather than smaller animals such as rabbits.

Dedrick and Casad (1999) list some more examples of N-V compounds, some of which have idiosyncratic meanings. All of the pairs provided by them appear in (68):
Yaqui N-V Compounds (Dedrick and Casad 1999: 161)

a. boó’o ‘road’ + bícá ‘see’ → boó’o-bícá ‘to wait for X’
   (‘watching the road’)
b. čoóni-m ‘hair-PL’ + poóna ‘pull’ → čón-poóna ‘be fighting’
   (lit. ‘pulling hairs’)
c. tékil ‘work’ + máka ‘give’ → tékil-máka ‘to commission, make responsible’
d. hiapsi ‘life/soul’ + temae ‘question’ → hiapsi-temáe ‘to repent’
e. táhi ‘fire’ + wéče ‘fall’ → táhi- wéče ‘have a fever’
   (lit. ‘fire-fall’)

It is not clear whether these N-V compound constructions are like (66) and (67) above and have incorporated nominals that cannot be modified by external elements such as adjectives, determiners, and the like. My analysis predicts that they should not be able to be so modified.

More common in this language is the use of incorporated nominals with derivational morphology, i.e. denominal verbs. The two most prevalent patterns are the denominal verbs of possession (AKA “verbless possessive sentences” or bahuvrihi constructions—Jelinek and Escalante 1988 and Jelinek 1998, respectively), and denominal verbs of making or using. The relevant morphemes are listed in (69) and (70):

Denominalizing morphemes in Yaqui I (Haugen in press b)

-ø HAVE
-e HAVE¹⁵
-te USE
-te MAKE

¹⁵ Dedrick and Casad posit a morpheme -ek as a denominalizing suffix to indicate the notion of HAVE, apparently because Sonora Yaqui does not allow for the full range of TAM marking on possessive denominal verbs that we see in Arizona Yaqui. That is, what Jelinek and Escalante identify as the perfective aspect marker Dedrick and Casad (1999) identify as an unrelated homophonous possessive morpheme.
Denominalizing morphemes in Yaqui II (Dedrick and Casad 1999)

-tu BECOME

'-u “intentive postposition” = ‘to go to perform some action on x”

-tua CAUSATIVE (e.g. ‘put on’)  

Yaqui possessive denominal verbs have a relatively substantial literature (see Jelinek and Escalante 1988, Escalante 1990, Martínez 1994, Jelinek 1998, Harley 2002, Jelinek 2003, Haugen in press b). In these constructions, nouns can be inflected with tense/aspect morphology in order to indicate the possession of that noun, as in (71):

(71) Peo kar-ek Peo has house(s)  

kar-ekan had houses  
kari-ne will have houses  
kari-su-kan used to have houses  
kari-pea feels like owning houses  
kari-vae is gonna have houses  
kari-maci should have houses  
kari-’eán should have (had) houses  

Previous analyses of these possessive denominal verbs have posited a zero morpheme for the verb of possession. In fact, Haugen (in press b), following Harley (2002), posits two null elements, a null possessive adpositional element (P_{HAVE}) and null stative light verb (v), and gives the following derivation for such verbs:

(72) aapo [ 'uka siali-k t₁ ] kar-ek  

3sg DET:ACC green:ACC t₁ house₁-PERF  

‘He has that green house’ (Jelinek’s translation: ‘That green house is his’)

(73)a. ... b. ...  

(Haugen in press: [33])

\[ \text{vP} \]

\[ \text{PP} \]

\[ \text{BE} \]

\[ \text{aapo} \]

\[ \text{P'} \]

\[ \text{DP} \]

\[ \text{P}_{\text{HAVE}} \]

\[ 'uka\ siali-k\ kari \]

\[ \text{vP} \]

\[ \text{PP} \]

\[ \text{BE} \]

\[ \text{aapo} \]

\[ \text{P'} \]

\[ \text{DP} \]

\[ [ \text{kar}_1 + \text{P}_{\text{HAVE}} ] \]

\[ 'uka\ siali-k\ t_1 \]
In the Appendix to Haugen (in press b), however, I suggest the possibility that there is actual diachronic morpho-phonological evidence for the possessive postpositional element $P_{\text{HAVE}}$—an element cognate with Yaqui instrumental case marker $-e$, which surfaces when these constructions appear in the habitual, in which case the nominal root is reduplicated:

(74)a. *Carmen a-hoara-po wi-wikich-e*  
*Carmen 3sg.POSS-home-in RED-bird-?
‘Carmen usually has birds in her home’

b. *Huan mo-mochik-e*  
*Huan RED-turtle-?
‘Huan usually has turtles’

c. *'aapo hiva ka-kava’-e*  
*3sg always RED-horse-?
‘He always has a horse’

Jelinek regards this suffix as a marker of “imperfective”; however, I suggest that it is in fact a postposition cognate with instrumental suffix $-e$, which is illustrated in (75) (from Dedrick & Casad 1999: 187 [26]):
(75) hunú-e matú-e hi’ib-oa mún-baki-m ’a-e bwá’a-bwása
    that-INST charcoal- INST eat-make bean-stew-PL it- INST RED-cook
    ‘She uses that charcoal for cooking; she cooks bean stew with it.’

If this is the case, then it is not obvious why this –e deletes in environments besides the
reduplicated one, cf. (71) above where all suffixes except perfective –k and past
perfective –ekan have an [i] between the nominal root and the TAM suffixes. However,
this is a larger problem within Yaqui morpho-phonology, since the same facts occur with
verbal roots with an ostensible –e ending, cf. vuite ‘run’:

(76) aapo vuite ‘s/he is running’       (Escalante 1990: Ch. 6 [15])
     vuite-k ‘s/he ran’
     vuite-kan ‘s/he had run’
     vuit-ine ‘s/he will run’
     vuit-su-kan ‘s/he used to run’
     vuit-ipea ‘s/he feels like running’
     vuiti-vae ‘s/he is going to run’
     vuiti-machi ‘s/he should run’
     vuiti-’eas ‘s/he should have run’
     vuiti-taite ‘s/he is starting to run’       (Maria Amarillas, p.c.)

The analysis of –e as a possessive marker is consistent with other intransitive verbs in
Yaqui, if we assume that some suffixes raise –e to –i, a phonological effect that is
completely general and not unattested in other languages.

Withholding opinion as to the desirability of separating verbal HAVE into an abstract
adposition incorporated into a light verb (as in Harley 2002), the analysis of –e as HAVE
in Yaqui is particularly attractive in light of the fact that a homophonous element is used
as a derivational morpheme in other very closely related languages. That is, we see the
same elements functioning as a derivational (denominalizing) suffix in both Guarijio and Tarahumara:\textsuperscript{16}

(77) Possessive \textit{\textipa{\textsc{e}}} in Guarijio\textsuperscript{17} (Miller 1996a: 149-50)

\begin{enumerate}
\item \textit{puhku} 'domestic animal' \rightarrow \textit{puhku-\textipa{\textsc{e}}} 'tener animal/have an animal'
\item \textit{atu} 'arma de fuego/gun' \rightarrow \textit{atu-\textipa{\textsc{e}}} or \textit{atu-\textipa{\textsc{e}}} 'tener arma/have a gun'
\item \textit{tehté} 'piedra/stone' \rightarrow \textit{tehté-\textipa{\textsc{e}}} 'tener piedra/have a stone'
\item \textit{wa\textipa{\textsc{kilá}}} 'camisola/shirt' \rightarrow \textit{wa\textipa{\textsc{kilá}-\textipa{\textsc{e}}} 'tener camisola/have a shirt'
\end{enumerate}

(78) Possessive \textit{\textipa{\textsc{e}}} in Western Tarahumara (Burgess 1984: 28)

a. \textit{gall} 'house' \rightarrow \textit{gal-e} 'have a house'

In addition to denominal verbs of possession, Yaqui also has denominal verbs of making and using. Some verbs of this nature appear to have a zero derivation, where habituality is indicated through reduplication of the nominal root:

(79) Verbs of using

\begin{enumerate}
\item \textit{aapo bo-bo\textipa{\textsc{c}}a-n} (Martínez 1994: 165)
\begin{tabular}{ccc}
3sg & RED & shoe-CONT
\end{tabular}
'S/he was using shoes'
\item \textit{chu-chu'\textipa{u}} (Martínez 1994: 163)
\begin{tabular}{ccc}
RED & dog
\end{tabular}
'have or use dogs'
\item \textit{aapo ka-k\textipa{\textsc{a}}ba\textipa{e}} (Martínez 1994: 170)
\begin{tabular}{ccc}
3sg & RED & horse
\end{tabular}
'He/she mounts horses'
\end{enumerate}

(80) Verbs of Creation

\begin{enumerate}
\item \textit{aapo ta-tahkae} (Martínez 1994: 169)
\begin{tabular}{ccc}
3sg & RED & tortilla
\end{tabular}
'S/he makes tortillas (habitually)'
\end{enumerate}

\textsuperscript{16} Although this suffix is also attested in these languages, this is not necessarily the most productive way of indicating possession in these languages. Many (if not most) Uto-Aztecan languages have multiple ways of indicating possession, and a minimal reconstruction for PUA possession would likely have something along the lines of a distinction between alienable and inalienable, with denominal verbs probably being included as the mechanism for one of these, if not both (Haugen in prep).

\textsuperscript{17} As in Yaqui, \textit{-\textipa{\textsc{e}}} can also serve an instrumental function in Guarijio (Miller 1996a: 286).
Verbs of spawning

a. hume kuchu-m si ka-kava
   DET-PL fish-PL very RED-egg
   ‘Those fish lay a lot of eggs’

b. hunu totoi husai-m ka-kava
   that chicken brown-PL RED-egg
   ‘That chicken lays brown eggs’

c. u tomato woi-m kaa bwasi-m taaka-k
   DET tomato.plant two-pl NEG ripe-pl fruit-PERF
   ‘That tomato plant has two fruits that are not ripe’

Some of these verbs are apparently relatively ambiguous between the possessive and creation (including birthing and spawning) readings.

In addition, some verbs of creation require the -te suffix, which is found in many other Uto-Aztecan languages:

Verbs derived with suffix: -te ‘make’

a. aapo kari-te
   3sg house-Make
   ‘S/he is building a house’

b. aapo kari-te-maci
   3sg house-Make-MODAL
   ‘S/he should build a house’

c. kari-te-’e!
   house-Make-IMP
   ‘Build a house!’

(83) aapo woim kari-te
    3sg two house-make
    ‘He is making two houses’

(84) Luis sami-te
    Luis adobe-Make
    ‘Luis is making adobes’

Like N-V compounding, these suffixes may not be fully productive, but rather seem to have a fixed set of particular nominal complements that can appear with them. Thus, many verbs of manufacture require a different verb, such as the free verb root hoa, and it is not ungrammatical to use this verb with some verbs that allow for incorporation into the -te light verb:
(85)a. *In akoo si kia burru-m ho-hoa* (Haugen in press b [41a])  
1sg.poss older.sister very delicious burro-pl RED-make  
‘My sister makes very delicious burritos’

b. *In akoo si kia tahkai-m ho-hoa* ([41b])  
1sg.poss older.sister very delicious tortilla-pl RED-make  
‘My sister makes very delicious tortillas’

c. *In akoo si kia tah-tahkae-ǝ* ([41c])  
1sg.poss older.sister very delicious RED-tortilla.MAKE  
‘My sister makes very delicious tortillas’

(86)a. *aapo kari-te.* (Maria Amarillas, p.c.)  
3sg house-MAKE  
‘he is building a house.’

b. *aapo kari-ta hoa.* (Maria Amarillas, p.c.)  
3sg house-ACC make  
‘he is building a house.’

c. *aapo káate* (Maria Amarillas, p.c.)  
3sg house.make  
‘he is building a house.’

All of the denominal verb constructions discussed so far allow for the stranding of modifiers, and thus instantiate what I have been calling “syntactic noun incorporation.”

Yaqui also has an additional denominal verb construction that appears to be an instantiation of classificatory noun incorporation. This is the construction that Jelinek (2003) refers to as the “USE-construction”, where a verb derived from one nominal root indicates that that element is comprised of some other entity, denoted by a nominal root appearing in the direct object complement. Some examples from Jelinek (1998: 217-18) are given below:

(87) *vempo [ 'uka kari-ta ] teopo-k*  
3pl DET:ACC house-ACC church-PERF  
‘They use that house as a church’ (lit. ‘They church-make that house’
(88) *inepo [ 'uka kuta-ta ] vo'onia-kan
   1sg DET:ACC stick-ACC cane-PAST.PERF
   ‘I had that stick as a cane’ (lit. ‘I caned that stick’)

(89) inepto [ 'uka 'ilii kuta-ta ] vo'onia-kan
   1sg DET:ACC little stick-ACC cane-PAST.PERF
   ‘I had that little stick as a cane’ (lit. ‘I caned that little stick’)

(90) *inepo [ 'uka kuta-ta ] [ 'ilii(-k) vo'onia-k ]
   1sg DET:ACC stick-ACC little cane-PAST.PERF
   ‘I use that stick as a little cane; I little-cane that stick’

As is characteristic of classificatory noun incorporation, there is no obvious syntactic slot from which the nominal root in the verbal position could have derived, since the direct object DP is headed by an overt, non-cognate nominal. However, it is crucial that the non-incorporated noun indicates some item that can fulfill the role denoted by the nominal appearing in the denominal verb. Thus, the direct object N is a hyponomous argument and this construction is an example of classificatory NI.

Finally, Yaqui allows for the option of including an incorporated pronominal on the verb, at least in the third person. This incorporated pronominal is marked for number:

(91) Inepo Hose-ta (aa)-vicha-k
    1sg Hose-ACC (3.SG.OBJ)-see-PERF
    ‘I saw Hose’

(92) Huan Hose-ta into Maria-ta (am)-vicha-k
    Huan Hose-ACC and Maria-ACC (3.PL.OBJ)-see-PERF
    ‘Huan saw Hose and Maria’

It is my claim here that these incorporated direct object arguments also act as classifiers, with the overt DP complement being a hyponomous object. This optional “object polysynthesis” can lead to full-blown object polysynthesis (i.e. O’odham-style obligatory
object agreement) when a new generation of speakers reanalyzes these morphemes as obligatory, thus setting the “object polysynthesis parameter” (see Chapter 9).

7.4.6. Corachol-Aztecan: Nahuatl

The various dialects of Nahuatl, Classical and Modern, show full-blown polysynthesis, meaning that transitive verbs in this language must be inflected for both subject and direct object. Object agreement can occur through noun incorporation or through an object “pronominal argument”. Because of the special status of polysynthesis with respect to other Uto-Aztecan languages, I will delay discussion of NI and denominal verbs in this language until Chapter 9, where I will argue that polysynthesis developed from classificatory NI in Pre-Aztec, a property inherited from PUA.

7.4.7. Denominal verbs across Uto-Aztecan

Most Uto-Aztecan languages have denominal verb morphology of some kind. Langacker (1977a) reconstructs the following morphemes for PUA:

(93) **Denominalizing Morphemes in Proto-Uto-Aztecan** (Langacker 1977a: 44-5)

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*-ka or *-kai</td>
<td>HAVE</td>
</tr>
<tr>
<td>*-tu</td>
<td>INCEPTIVE</td>
</tr>
<tr>
<td>*-ta</td>
<td>MAKE</td>
</tr>
</tbody>
</table>

Haugen (in prep) provides a recent survey of denominal verbs across Uto-Aztecan, and confirms these reconstructions. I also argue that there was probably an additional possessive marker, and the two probably were differentiated along the lines of an alienable/inalienable distinction. Because of the wide distribution of morphemes across the family, we probably also reconstruct denominal verbalizers (v’s) meaning USE,
MARRY, GET (or GATHER), and REMOVE. In addition, it seems to be the case that these derived verb constructions typically allow for modifier stranding and hyponomous objects.

7.5. Conclusion

In this chapter I have argued that the distinction proposed for noun incorporation and denominal verb constructions is a false one. Whereas noun incorporation is an option employed by speakers for certain verbs, other verbal elements (i.e. denominalizing verbal affixes) require this incorporation. However, in many instances paraphrases are available, with full verb roots, to express the meanings typically served by these affixes (e.g. making or having).

I argued that the evidence from noun incorporation in Uto-Aztecans suggests that we must revise Mithun (1984)'s implicational hierarchy of NI types, because we see Type 4 (classificatory) noun incorporation in languages without Type 3 NI. I argued further that Type 3 NI, or object polysynthesis, is in fact an instantiation of Type 4 NI.

Finally, a survey of NI and denominal verb constructions across Uto-Aztecans was presented. While all Uto-Aztecans languages have denominal verbs, and share many cognate denominal verb affixes, NI differs in these languages. Most languages illustrate hyponomous arguments with denominal verbs at least, and at least some have hyponomous arguments with NI as well (e.g. Nahuatl, Hopi). Some languages have optional object prefixes (Cupeño, Yaqui), and some require object prefixes (Tohono O'odham, Nahuatl). Hopi is interesting because it has a mixed system wherein only a
handful of verbs require object prefixes, thus showing an intermediate stage between optional and obligatory object prefixes. The grammaticalization of NI across the family will be the topic of Chapter 9.
CHAPTER 8
ON THE SYNTAX OF NOUN INCORPORATION: INCORPORATION AND CONFLATION, MOVE AND MERGE

8.1. Introduction

In Chapter 6 I focused on the claims of Mithun (1986) that suggest that noun incorporation (NI) should be viewed as a fundamentally different process than denominal verb formation. From Mithun's perspective, noun incorporation is a morphological process akin to compounding, whereas denominal verb formation is a derivational process (also morphological) where a noun stem is added to some category-changing suffix. Because both denominal verbs and verbs with incorporating nouns in Hopi (and other Uto-Aztecan languages) share similar properties (modifier-stranding and the classificatory function) in the usual case, I argued that the analyses of the two processes should be linked under a syntactic account of word-formation. In addition, after critiquing Rosen's Lexicalist account of NI, I alluded to the desirability of presenting an analysis in which noun incorporation (and by extension, denominal verb formation) occur in pure syntax, as has also been argued by Sadock (1980, 1986) and subsequent researchers, and as is consistent with the Pervasive Syntax Perspective (PSP) adopted in Chapter 3.

In Chapter 7 I argued that, descriptively at least, we can make a distinction between several varieties of NI attested in the Uto-Aztecan languages. These are N-V compounding (or "classical NI"), syntactic NI, object polysynthesis, and classificatory
NI. However, these terms were only intended to represent heuristic categories for use in cross-linguistic description. In this chapter I will argue that each of these four categories can in fact be derived through a single syntactic mechanism: incorporation, the product of Move, construed as head movement (Baker 1988), when the incorporated noun functions as the direct object of the verb. A second syntactic mechanism, Merge, can derive N-V compounding constructions to yield non-object functions for a nominal root attached to a verb. This is conflation (Hale and Keyser 2002). Incorporation will account for each of the types of NI that we see when the incorporated nominal appears as the direct object of the verb (sections 8.3-8.5). Conflation will account for those cases of word-formation that involve manner elements such as the instrumental prefixes found in Numic (section 8.6).

In the sections that follow I will introduce the theoretical apparatus that executes each of these processes in syntax from the non-Lexicalist, Distributed Morphology perspective. Here I will follow recent discussion of these issues that goes beyond the phenomena traditionally related to NI, but also have been related to denominal verbs and compounding more generally.

8.2. Theoretical background: Incorporation and Conflation

In section 8.2.1 I review the prominent proposals for deriving incorporation through head movement—Baker (1988) and Hale and Keyser (1993), and in section 8.2.2 I discuss Hale and Keyser (2002)'s more recent approach to these issues, conflation, and I argue that this notion actually has a much more limited applicability than they themselves
suggest. I argue that although their wholesale abandonment of the head-movement account of denominal verbs was premature, their notion of conflation as a concomitant of Merge does have application in non-object NL.

8.2.1. Hale and Keyser (1993): Denominal verbs and incorporation

In the strongest version of the PSP to date, Hale and Keyser (1993) claim that all denominal verbs are formed via head movement, in the sense of Baker (1988). Further, they also claim that denominal verb formation is formed by such movement by definition. They state that "it is a fundamental assumption of our account that English verbs like shelf, and other location verbs, are ‘denominal’ precisely in the sense that they are derived by head movement" (pp. 56-7). In their 1993 framework, this involves the notion of “Lexical Relational Structure” (LRS, referred to in subsequent literature as “L-Syntax”).

In this framework, the LRS of a morpheme determines its argument structure. LRS is not equivalent to “D-Structure” (Chomsky 1965), although it determines the D-Structure representation of a verb, and it is governed by the same principles that constrain syntax proper (e.g. the Head Movement Constraint of Travis 1984). In the 1993 framework, this is “lexical” in the sense that it happens out of syntax proper, although Hale and Keyser do conclude by saying that it would be better to not have the same processes occurring in “lexical syntax” and “syntax syntax” (my phrasing), since, for example, unergative and unaccusative verbs do behave differently in syntax. An example of such differences is
that "simple transitives" fail to undergo causative or middle alternations, while "ergatives" do, as contrasted in the following examples (Hale and Keyser 1993: 89):

(1) a. We smeared mud on the wall.
   b. ?Mud smeared on the wall.

(2) a. The pigs splashed mud on the wall.
   b. Mud splashed on the wall.

Hale and Keyser (1993) distinguish these two classes by giving the latter a "manner component" in its LRS, while the former has to get the manner in syntax (e.g. in vP).

An example of Hale and Keyser (1993)'s derivation of the English denominal verb *to shelve* is given in (4). The lexical relational structure of the verb *to shelve* corresponds with the lexical relational structure of a ditransitive verb like *put*, shown in (3):

(3) ditransitive verb: *put*

```
(3) ditransitive verb: put

\[ \text{V'} \]
\[ \text{V} \]
\[ \text{VP} \]
\[ \text{V} \]
\[ \text{NP} \]
\[ \text{V'} \]
put (her books) V PP
\[ \text{t} \]
\[ \text{P} \]
\[ \text{NP} \]
(on the shelf)
```
(4) denominal verb: shelve

```
V'          V
  |    VP
shelf     NP    V'
               |    PP
                |    V
               (her books)  P    NP
                    |    t
```

Under Hale and Keyser's (1993) analysis, the locational denominal verb projects the same argument structure as the verb put, but the former is formed through incorporation of the lexical head (here, shelf) into an abstract, null preposition; incorporation of this complex into the lower V; and, finally, the incorporation of this complex into the higher V, which corresponds to the more recent theoretical construct "little v" (Kratzer 1996).

More recently, Hale and Keyser (2002) have abandoned this approach because of what I will call "the hyponomous object problem": sentences such as John shelved the books on the radiator indicate that these verbs, supposedly arising through incorporation, can in fact take direct objects that are not "cognate"—i.e. that are not derived from the same root as that which creates the denominal verb. In fact, this is the same issue that Mithun and Rosen have raised with respect to "Classificatory NI", and, after further exposition, I will propose the same solution to this problem for each of these cases in section 8.3 below.
Along with Hale and Keyser (2002), we may differentiate *hyponomous* objects from *cognate* objects. In the case of cognate objects, there is root identity between the verb and the direct object (or object of the PP). Hyponomous objects are not root identical, but the root must denote a subset of the set of entities denoted by the root in the verb:

(5) **Cognate Objects—w/ root identity**
   a. He slept the sleep of the just.
   b. He laughed his last laugh.

(6) **Hyponomous Objects—object is not root identical**
   a. He danced a jig.
   b. He bagged the potatoes in a gunnysack.

Note that in the case of cognate objects in English, it seems that the object must co-occur with modifiers or other restrictors.¹ This is not true of hyponomous objects (cf. the contrast in 7 and 8):

(7)a. *He slept a sleep.
    b. ?He laughed a laugh.

(8)a. *?He danced a dance.
    b. He danced a deer dance.
    c. He danced a jig.

According to Hale and Keyser (2002), there are syntactic differences between the two, as illustrated by the (admittedly subtle) semantic differences of the two classes. Some verbs permit only true cognate objects, where no hyponomous objects are allowed:

(9)a. *She slept her last nap.
    b. *He laughed a surreptitious giggle.

Also, cognate objects do not seem to allow pronoun replacement under conjunction (ex. 10), whereas hyponomous object verbs accept such readily (ex. 11):

¹ I owe this observation to Heidi Harley.
(10)a. *John slept the sleep of the just and Bill slept it too.
   b. *Robin laughed the laughs of the Rat Pack, and Jonathan laughed them too.

(11)a. Robin sang the songs of the 60s and Jonathan sang them too.
   b. John danced a tango and Bill danced one too.
   c. John danced the tango and Bill danced it too.

These facts raise the question of why such differences should arise. One possibility mentioned by Hale and Keyser (2002), which I will defend in detail below, is that strict cognate object verbs arise through incorporation, where both the head and trace of a movement operation are spelled out. An example of such a derivation, using Hale and Keyser (2002)'s proposed structure, is given in (12):

(12)

\[
\begin{array}{c}
V \\
\hspace{1cm} V \\
\hspace{2cm} N \\
\hspace{3cm} \text{sleep}
\end{array}
\begin{array}{c}
\hspace{1cm} V \\
\hspace{2cm} \text{the}
\end{array}
\begin{array}{c}
\hspace{1cm} D \\
\hspace{2cm} \text{NP}
\end{array}
\begin{array}{c}
\hspace{1cm} \text{sleep of the just}
\end{array}
\]

It should be pointed out that the structure in (12) presents a problem for Hale and Keyser (2002)'s assumptions about movement, and that is that the head of NP, N, should not be able to incorporate out of DP, because of the Empty Category Principle and the Head Movement Constraint. Hale and Keyser discuss two options for getting around the problem of the ECP. First, it is possible that DP is not a barrier to extraction. Second, it might be the case that the ECP does not apply here since there is no actual empty category: the trace of the movement is actually spelled out.
However, since a head (D) intervenes between the N and V, the Head Movement Constraint remains violated. On this issue Hale and Keyser note that “this is a general problem, and determiner stranding, if it actually occurs, flies in the face of it (though it does not do so in Baker’s (1988) analysis of the constructions)” (p.73).

Although Hale and Keyser waver here as to whether or not they actually believe in determiner stranding, I think that the empirical evidence, much of it reviewed in Chapters 6 and 7 with respect to NI in Uto-Aztecan, is clear on this point, and that such facts must be somehow accommodated in whatever syntactic theory is to be adopted. The problem raised here is particular to the theory of syntax that includes DP as the complement to V, and evaporates immediately if we suppose that DP is not in fact a barrier to the of extraction of N.

In the interval between their work in 1993 and their monograph in 2002 Hale and Keyser dispose of their core idea that denominal verb formation is identical to noun incorporation, at least in their analysis of English. The solution that they propose to get around the problem of hyponomous and cognate objects is the process that they term conflation, which we will discuss in section 8.2.2, and revise in section 8.6. However, although they ultimately abandon their syntactic approach to denominal verbs, I will argue presently that this abandonment was premature (cf. also Harley 2002, in press), and that their notion of conflation has a much more limited application than they suggest. Since the same problems exist for noun incorporation (i.e. “Classifier NI” is an instantiation of incorporation with hyponomous arguments) as for denominal verb formation, I think that the two are in fact as intimately linked as Hale and Keyser
proposed in 1993. What is needed is a way to account for hyponomous arguments, and this is done through the DM notion of the Late Insertion of lexical material.

Before turning to this discussion, however, I should point out that Hale and Keyser (2002) do not eliminate head-movement incorporation from their theory of English altogether. In Hale and Keyser's earlier work, denominal verb formation occurred in L-syntax, but under the tenets of DM, the essentials of which they had adopted by 2002, there is no independent notion of "lexical syntax"—all syntax (including word-formation) is syntax proper. While Hale and Keyser (2002) find the incorporation analysis for English locational denominal verbs to be untenable, they do maintain this analysis for English deadjectival verbs (e.g. 'The sky reddened'), as in (13):

\[
(13) \quad \begin{array}{c}
V_1 \\
\downarrow \\
V_2 \\
\downarrow \quad \downarrow \\
DP \\
the sky \\
\downarrow \\
A \\
\downarrow \quad \downarrow \\
red \\
\downarrow \\
en
\end{array}
\]

A similar analysis can also be posited for deadjectival verbs in Uto-Aztecan, as Hale and Keyser show with examples from Tohono O'odham (2002: 130-2).

8.2.2. Hale and Keyser (2002): Denominal verbs and conflation

In order to bypass the problem posed by hyponomous arguments, Hale and Keyser (2002) make a distinction between two separate processes: incorporation, involving the
operation Move, and conflation, involving the operation Merge. Both notions involve the formation of a complex predicate. It is my position that the latter is intimately related to the notion of compounding.

For Hale and Keyser (2002), incorporation of the kind that we have seen throughout Uto-Aztecan can be distinguished from conflation, in that the latter never leaves "a residue" of stranded modifiers. As discussed in earlier chapters, the modificational "residue" in the following Hopi sentences is argued to be the result of movement of the head noun out of NP and into v:

(14)a umu-na kaway-mu-y kuk-hep-ma 
   2pl-father horse-PL-GEN track-seek-GO:PERF
   'Y'all's father has been to search for horses' tracks'

b. Nu' pu-t ki-'yta
   I that-ACC house-HAVE
   'I have that as a home'

Hale and Keyser posit a syntactic source for the stranded modifiers in such denominal verbs, much like what Sadock (1980, 1986) had argued for Greenlandic, and which was argued against by Rosen (1989). With conflation, supposedly, such a residue is never left.

A derivation of an incorporation construction, with the denominal possessive verb of (14b), is shown in (15):
According to the theory developed by Hale and Keyser (2002), the crucial difference between incorporation and conflation is the notion of strict complementation, defined in (16):

(16) **Strict Complementation** (Hale and Keyser 2002: 59)
A head X is the strict complement of a head Y iff Y is in a mutual c-command (i.e. sister) relation with the maximal categorial projection of X.

This notion places a limit on what can “confl ate”—a head can only conflate into another head if it is in a phrase that is a sister to the head into which it is conflating; in other words, a head Z cannot conflate into head X if there is an intervening maximal projection YP:

(17) 

For a more concrete example, consider (18):
Here, the conflation of N into V via P in the derivation of *corral* is ok, because NP is sister to P, and PP in turn is sister to V: in each case of conflation the requirement of strict complementation is met. What would not be allowed is the direct conflation of N into V without first conflating with P. This appears to restate the Head Movement Constraint, and this account is in fact very similar to the incorporation account, except for the additional notion of "strict complementation". This move, however, is problematic, since what matters in evaluating the process of movement under this metric is solely the issue of whether the complement of a V is DP or NP. Under this account, conflation is ok with NP as a complement to V, but not with DP, since with NP embedded under DP there is an intervening head, and N should not be able to conflate (unless it conflates into D first).

Note also that Hale and Keyser define strict complementation so that the requirement holds over the maximal projection of a head X. One result of this requirement is that specifiers should not conflate—i.e. local c-command and government are not enough for conflation: a conflating X has to be sister to N, V etc., not DP, TP, etc.

Consider the result of this for (19), the tree for the sentence in (14)b:
According to Hale and Keyser,

While incorporation of *kii- ‘house’* is possible, giving the denominal possessive verb *ki-’y- ‘have house’* and stranding the determiner *pu- (ultimately the accusative form *pu-t)* ‘that’, conflation is impossible in this case because N is not a sister of the matrix verb, as would be required of conflation in conformity with the strict complementation requirement. (p.60)

Thus, for Hale and Keyser the crucial distinction between conflation and incorporation is the level of embeddedness of the affected N: if it is under DP it can incorporate but not conflate with V. If an N was instead embedded under an NP complement to V, then it should be able to conflate, since strict complementation would be met.

However, if there was something else in NP (say, an adjective), this account does not explain why there would be no “residue”. The 2002 approach also does not rule out conflation of N with D prior to conflation with V, parallel with the locational verbs (with conflation from NP to PP, then to VP) above.
Thus, Hale and Keyser’s revised approach, where conflation is basically incorporation with an attempt to further specify the possible restrictions on movement, does not actually produce the results that Hale and Keyser desire. Either conflation is not a separate notion from incorporation, or conflation is to be defined in some other way. I suggest the latter. In my view, conflation is the “compounding” of a root with another root, or crucially, the compounding of a root with a functional head. This definition of conflation will be discussed in more detail in section 8.4.

In the next section I will show how the syntactic incorporation account can be maintained for Uto-Aztecan denominal verbs and other NI constructions with full DP hyponomous object complements.

8.3. Late Insertion as the solution to the hyponomous object problem

To briefly sum up the theoretical discussion so far, Hale and Keyser (2002) abandon their (1993) head-movement account of English denominal verbs, wherein denominal verbs were formed syntactically via incorporation (in the sense of Baker 1988), largely due to the hyponomous object problem: in examples such as (6), where the direct object is non-cognate with the denominal verb, there is no obvious source for the lexical material that is contained in the verb.

Similar facts obtain in some languages with overt NI, in what Mithun (1984) terms “classificatory noun incorporation”, which involve those cases of NI where a verb may incorporate a noun but still take a non-cognate but semantically related direct object.

---

2 Hale and Keyser themselves suggest that this is a possibility (2002: 69).
Such facts have led some scholars (e.g. Rosen 1989; Mithun 1986, 1999) to reject the PSP and propose Lexicalist theories of noun incorporation, contra Sadock (1980, 1986), Baker (1988), and Hale and Keyser (1993). In this section I propose a reconciliation of these facts for a purely syntactic approach to NI by presenting an analysis of NI that utilizes a Late Insertion theory of “traces” (in the Minimalist program, copies) of movement.

It should be recalled from Chapter 3 that, by hypothesis, the structures actually manipulated by syntax are bundles of features. Morpho-phonological material only enters into a given derivation post-syntactically, at Morphological Structure, when “lexical” material is inserted into the conglomerations of features created in syntax. This is the idea of Late Insertion. We can resolve the hyponymous argument problem by allowing for the insertion of hyponymous arguments into the position of the lower copy (“trace”) left over after movement. Although I assume that this theoretical move has general applicability to zero-derived denominal verbs of the English kind, I will focus my attention on the overt morphology involved in NI and denominal verbs in Uto-Aztecan, reviewed at length in Chapter 7.

Consider again the facts displayed by Classificatory NI with denominal verbs in Hopi (20)b, Yaqui (21), and other Uto-Aztecan languages (22), all of which show instantiations of hyponymous objects with incorporated nominals:

(20) **Hopi Denominal Verb with Cognate Object**
    a. *Hak* yóypu-t aayga-t aay-an-numa.  (K. Hill 2003: 239 [131])
    who cracked-ACC rattle-ACC rattle-CAUS-CIRC

    ‘Someone’s going around shaking a cracked rattle.’
b. **Hopi Classificatory NI: NI with a Hyponomous Object**

*Pam* tsiili-t *nakwa*-y-ta-ngwu.  
(he *chile*-ACC *feather.worn.on.head*-POSS-DUR-HAB)

‘He (the Hehey’a kachina) wears chili pepper (“chile as a feather”) on his head.’

(lit. ‘He feather.worn.on.head-has the chili pepper’)

(21) **Yaqui Denominal verbs with a Hyponomous Object**

*u*ka *ili chu’u-ta ne=vuk-ek*  
(det.ACC little *dog*-ACC 1.sg.SUBJ=pet-PERF)

‘That little dog is my pet’ (lit. ‘I pet-have that little dog.’)

(22) **Hyponomous objects with denominal verbs in other Uto-Aztecan Languages**

a. **Tümtpisa Shoshone**  
*Nüümü so’oppüh putish pungkupaimmippihantü*  
(use(exc) many *burro*-PET-HAVE-PAST)

‘We used to have many burro pets’

b. **Gosiute Shoshone**  
*Isapaippih sukka ponaiha taipai*  
(Coyote that-OBJ *Mouse*-OBJ brother-HAVE)

‘Coyote has Mouse for a younger brother’

c. **Cupeño**  
*tüku=’ep ne-’ásh-lyu awá-l-i*  
(yesterday=r 1s-pet-VB *dog*-NPN-O)

‘Yesterday I had a dog’

The Hopi data in (20) show NI with two different kinds of direct objects. In (20)a we see a cognate direct object, where there is root identity between the object and the verb (i.e. *aaya* ‘rattle’). In (20)b we see a hyponomous direct object where the direct object noun is a subset of the class of nouns denoted by the incorporated noun. Following the Copy Theory of Chomsky (1995), we can suppose that the overt direct object in (20)a is the phonological spell-out of the lower copy of the moved noun, as in the following derivation (the final derivation in 23b indicates the Accusative Case-Marking omitted...
from 23a; Tense, Aspect and Agreement nodes are abbreviated as “IP” for INFL-P, and specifiers, including the subject position of Spec of vP, are ignored):

(23) *H*ak yóypu-*t* aava-*t* aav-*an*-numa.

\[ \text{who cracked-ACC rattle-ACC rattle-CAUS-CIRC} \]

‘Someone’s going around shaking a cracked rattle.’

The claim here is that the spell-out of the lower copy in constructions like this are equivalent to the spell-out of the lower copy in such English examples as (5) above. The difference between Hopi and English is that Hopi allows for null-head modifiers and this lower copy is optional, but it is obligatory in English.

Further cross-linguistic support for this movement account comes from languages like Hindi, which also has obligatory spell-out of the lower copy in certain denominal verbs, such as *khaana khaa-* ‘food-eating’ and *gaanaa gaa-* ‘song-singing’ (Klaiman 1990: 331). The puzzle here lies in how to provide a syntactic account for (20)b, where there is no such overt trace of the root that has undergone head movement, and in fact there is a non-cognate, hyponomous argument in that argument position.

My solution to the hyponomous argument problem is that the lower copy may be spelled out with root material that meets more specific featural specifications than the
higher copy. Some languages allow for the grammaticalization of classifiers to be 
inserted into the highest node of a chain formed by head movement. In some languages 
this classifier system can be quite complex, as for instance with Caddo (Caddoan) 
(Mithun 1984, Melnar 2004). In Caddo, there is a closed class of a number of 
incorporating nominal elements that Melnar (2004) calls "patientive". This class of roots 
"includes elements that indicate body parts (e.g. k'ánt- 'head', ńčah- 'eye'), body 
products (e.g. bak(a)- 'word', kák ušt- 'saliva'), cultural products (e.g. nisah- 'house', 
t' ánk- 'pipe), and natural phenomena (e.g. háwt- 'wind', wadát- 'earth')" (p. 170). 
According to Melnar, 

as part of the derivational machinery of Caddo, incorporated nouns and other 
patientive prefixes modify the meaning of the verb by narrowing the sense of its 
patient. In this respect, all patientive elements have a classificatory function. As 
classifiers, patientives necessarily indicate a class of items and do not specify any 
particular class member (p. 171). 

However, "each patientive term has a default interpretation that is consistent with the 
prototypical or stereotypical member of its class. This default interpretation is the 
'elsewhere' meaning that comes into play when no other more specific meaning is 
contextually available" (p. 171). For example, the default interpretation for kan- 'liquid' 
is 'water'. I assume that each of these classes of roots correlate to some feature (or 
bundle of features) made available to the syntax by Universal Grammar. In Uto-Aztecan, 
the typical feature of relevance in incorporated pronominals is often simply [+3rd person],
or the equivalent. In Caddo, however, there must be more specification of features, for which these classifiers are inserted at MS.

Harley and Ritter (2002) present a feature geometry for the instantiation of pronouns in Universal Grammar. Their proposal is given in (24):

(24) Morphosyntactic feature geometry for pronouns (Harley and Ritter 2002: 486 [6])

As they suggest, this proposal could be falsified in one of two ways. First, some empirical facts could suggest a revision of the actual hierarchy of features, but still maintain that a feature-geometric approach is ideal for the expression of features relevant to syntax. However, alternatively, it could be argued that the geometric approach itself is untenable by showing that any geometry would be inadequate to account for a certain paradigm. For example, a form in a paradigm might show syncretisms that demonstrate that some node would need to be dominated by the Participant and Individuation nodes simultaneously (for instance, an inclusive that collapsed sometimes with 1st person singular forms and sometimes 3rd person plural forms) (p. 518).
Facts such as this would indicate that “the geometric approach to morphological features is fundamentally flawed” (p. 518).

Agreeing with Harley and Ritter that this second development would be highly unlikely, I assume that something like (24) is the appropriate way to represent the features available for syntax. Extrapolating from the pronominal domain to the area of other referential elements (i.e. other nominals), I assume that the features relevant for classifiers are subsumed under the Class node. These can include features like those Harley and Ritter include in their hierarchy ([+/ animate], grammatical gender, etc.), but also features that are typically referred to by classifiers cross-linguistically, such as reference to shape, etc.

In the case of Caddo, the Class node would have to branch into at least four different sub-classes to indicate Melnar’s differentiation of classifiers indicating body parts, body products, cultural products, and natural phenomena. Rood (1996) gives a hierarchy of semantic features for the closely-related language Wichita:

(25) **Semantic selectional features that subclassify Wichita nouns** (Rood 1996: 592)

```
  nouns
  /   \
 +count -count
   /   \
+collective -collective liquid dry mass
      /   \
 animate other activity
```

I assume that the ordering of these features is a language-particular arrangement of probably universal features, since I assume that for the purposes of syntax the realization of the lexical material inserted at Spell-Out is not relevant; i.e. it is not relevant what the
particular roots actually mean. For example, the notion of 'saliva' (Caddo *kák'uši*) does not play a role in the syntax per se (Harley and Noyer 2000), although it can be inserted as a lexical item spelling-out some higher order feature, such as [-animate], perhaps in conjunction with other features.

Returning to Uto-Aztecan, where the crucial feature is typically [+3rd person] or its equivalent, the idea is that when this feature is inserted into the syntactic slot of the complement to V and incorporates, it is possible to spell this out with more than one root capable of realizing this feature. A simple, hypothetical derivation of noun incorporation with a third person singular form is given in (26), where (26)b is the result of incorporation:

(26)a. . . . . b. . . . .

\[
\begin{array}{c}
\text{VP} \\
\text{V} & \text{NP} \\
\triangle & [+3^\text{rd}, +\text{sg}] \\
\end{array}
\hspace{1cm}
\begin{array}{c}
\text{VP} \\
\text{V} & \text{NP} \\
\triangle & [+3^\text{rd}, +\text{sg}] \\
\end{array}
\]

After head movement, the incorporated head must be spelled out in its highest syntactic position—the head of the chain that it forms with its lower copies, "traces" in earlier theories. In the very simple derivation in (26), there are two ways to spell out the [+3rd] and [+sg] features inserted into this derivation. In the case of languages with fully productive NI, any root that could spell out a [+3rd] and [+sg] features could be inserted into the incorporated nominal position; i.e. almost any nominal root could be incorporated, along with the third person singular pronominal. This would be a language
like Nahuatl. A second possibility would be a language that has grammaticalized an incorporated classifier for this position. I suggest here that languages with obligatory agreement markers are such languages, e.g. Tohono O'odham. In this case, the pronominal object argument spells out the person and number features, a development along the path of grammaticalization from full-blown NI (with incorporation of roots and pronominal elements) to pronominal arguments (to the exclusion of nominal roots) (see Chapter 9).

One can imagine highly complex bundles of features, however, and it is in these languages, like Caddo, that further specification of the feature bundles would be required. What is important here is the essential idea behind (26). The link between the incorporated nominal and its hyponomous argument is created via head movement. The lower copy (i.e. the “direct object”) has minimally the same features as the upper copy (i.e. the incorporated element), but it can also be spelled out with roots specifying even more specific features. It appears that the lower “copy” may be spelled out only in the event that it appears in a phrase which gives further specification delimiting the kind of head in question, as per examples (7) and (8) above. Typically, hyponomous arguments are inserted to denote membership in the semantic class denoted by the incorporating head. However, a hyponomous construal can be coerced by inserting practically anything into the lower position. An example this is the Hopi sentence in (20)b, where the incorporated noun nakwa ‘feather.worn.on.the.head’ classifies its direct object: tsili ‘chili pepper’, as an object that acts as a feather to be worn on the head.

Because the lower copy remains coindexed with the upper copy, the Late Insertion of
a different set of roots into the lower copy does not constitute a violation of the Projection Principle—the material that spells out the lower copy must minimally satisfy that set of features instantiated in the upper copy. The restriction on spelling out the lower copy of an inserted root only in those cases where further information is specified is taken to be a pragmatic condition, as encoded, for example, in Grice's Maxim of Quantity—cf. English *sleep sleep, *dream dreams, etc.

Hale and Keyser themselves have attempted a similar proposal to account for hyponomous objects in English denominal verbs (Hale and Keyser 1997), but they subsequently abandoned it (Hale and Keyser 2002). The Hale and Keyser (1997) proposal differs from the one being advocated here, however, in that their proposal involves “de-linking” the head of a movement chain from its lower links, through the deletion of the indices on the head of the movement chain and its traces. Consider the sentences in (27) (Hale and Keyser 1997: 42 [21]):

(27)a. She shelved the book.

b. She shelved the book on the top shelf.

Hale and Keyser (1997) analyze the verb shelve in (27)a just as above in (4), where the denominal verb is formed via incorporation into an abstract preposition and then into v. In (27)b, however, they argue that “the verb in [27b] has entirely lost its referential character. Here we know that a shelf is involved only because of the shelf appearing in the PP. The sentence is as fully grammatical with windowsill, desk, mantle, or sawhorse in place of shelf” (1997: 43). The (denominal) verb itself only denotes a particular kind of putting: that prototypically involved in stacking items such that “the entity denoted by
the direct object must be in a condition which will permit it to adopt a particular stance (typically upright) in its destined location” (p. 43).

They then go on to state that

if the verb of [27b] is derived by means of incorporation, as is the verb in [27a],
then we must assume that our proposed index-deletion process has applied,
effectively removing the referential connection between the incorporated nominal and its original position. . . [s]yntactically, this results in the circumstance that the syntactic heads involved in the original movement chain. . . no longer represent a chain, since the traces are effectively gone, due to index-deletion. (p. 43)

Hale and Keyser (2002) dismiss this attempt at reconciling the facts of hyponomous arguments with their derivation of denominal verbs via incorporation, saying that they would not further reconsider their 1997 analysis,

according to which hyponomous arguments and cognate objects alike were the result of ‘reinsertion’ into the trace position created by conflation [sic]. This is an impossible notion, not only for general theoretical reasons. It is also impossible, within the theory considered in [Hale and Keyser 2002], because conflation is not head movement and hence does not leave a trace in any conventional sense (p. 88).

In my proposal, the operation in question is head-movement, so there is a “trace” left behind after movement, but this “trace” is a copy of the relevant features. Second, no “de-linking” of indices is necessary, because with Late Insertion the spelling-out of morphological material involves minimally the same features. The crucial restriction is
that the head of the chain must have the less specific features spelled out; in this way, the hyponomous object relation is derived via head movement.

If there is some theory-internal reason why the Late Insertion of material cannot occur in the manner suggested here then the PSP is in much jeopardy. Not even withstanding the English-based objections of such critics as Kiparsky (1997), this approach must also account for the obviously syntactic relations that hold between incorporated nominals that serve a classificatory function and hyponomous arguments, which are so robustly displayed by the Uto-Aztecan languages.

One thing that should be stressed here is that the relations created between a classifying element and its hyponomous object cannot be limited to only formal feature-geometric relations. There must be some interaction between the syntactic licensing of roots and the Encyclopedic knowledge used in interpreting those roots in the syntactic contexts in which they appear, since there can be novel usages of lexical items as classifiers (e.g. tsiili ‘chili pepper’ for nakwa ‘feather worn on the head’ in Hopi), and in other examples.

In syntax, the things being manipulated are simply bundles of features, thus there is nothing inconsistent about spelling out two different lexical items that share the relevant features. What is necessary is an apparent stipulation that roots with more specific features than those given by the syntax be able to be inserted into lower positions of a chain, rather than at the head. The approach advocated here allows for parametric variation, in that some languages may require the upper copy to be the spell-out of the highest member of q semantic hierarchy, e.g. in those languages which have obligatory
classifiers serving as incorporated nouns on the verb (e.g. Caddo). In addition, not all languages allow for the insertion of hyponomous arguments, e.g. West Greenlandic and Southern Tiwa. As Rosen (1989) points out, however, this is a problem for both the Lexicalist and non-Lexicalist syntactic accounts of NI. I suggest here that this restriction is a language-specific parameterization.

To conclude this section, I have argued that classificatory NI is reconcilable with straight-up syntactic NI, in that hyponomous objects are allowable under the Late Insertion theory of morphosyntax. In the next section we examine cases of non-DP complements to v in Uto-Aztecan NI constructions.

8.4. Non-DP complements to V: N-V compounding revisited

Ironically, the syntactic approach to NI advocated in section 8.3 raises a problematic issue for what has hitherto been the most well-understood (and in Lexicalist theories, easy to explain) kind of NI, that which has been recognized since Sapir (1911). This is the “classical NI” that I have termed N-V compounding. Recall from Chapters 6 and 7 that it appears that this kind of NI must be separated from syntactic NI, because, descriptively, it “detransitivizes” a verb and thus obviates the ability of that verb to take a direct object, such as the “null-headed” direct objects identified via “stranded” modifiers (Rosen 1989). This is seen in such Yaqui examples as the following (repeated Chapter 7):
(28)a. *aapo maaso-ta peu-ta-*k  
   3sg deer-ACC butcher-TRAN-PERF  
   ‘He butchered a deer’

b.  *aapo maaso-*peu-te-n  
   3sg deer-butcher-INTR-PAST  
   ‘He was deer butchering’

c.  *aapo bwe’uu-k maaso-peu-te-n  
   3sg big-ACC deer-butcher-INT-PAST  
   [‘He was [big deer]-butchering’] or [‘He was deer-butchering a big one’]

The question raised by data such as this is how a theory coming from the PSP should account for compounding: is it lexical, or is it syntactic? Many scholars take evidence like that in (28) as definitive proof that the Lexicon is generative (e.g. Mithun and Corbett 1999), although some have wavered and allowed at least for the possibility that in principle syntactic NI could co-exist with lexical NI (Rosen 1989); Baker (1995) himself has recently weakened his own theory to allow for some lexical NI co-existing with syntactic NI.

In section 8.4.1 I discuss my theoretical account of N-V compounding in sentential syntax, and in 8.4.2 I show that the suppositions propounded here have cross-linguistic application that exceeds the bounds of comparative Uto-Aztecan morphosyntax.

8.4.1. N as complement to V: A formal account of N-V compounding

My analysis of N-V compounding, framed within the PSP outlined in Chapter 3, may seem at first blush to be a bit radical. What I will suggest is that verbs in N-V compounding constructions are not actually “intransitive”, as they appear on the surface. Rather, they sub-categorize, perhaps optionally, for a bare nominal complement rather
than a full DP. Since Case is actually a requirement of DP rather than N, the nominal in this configuration cannot receive Case and is thus forced to incorporate into V, as is outlined by Baker (1988). It is in the complement to V position that the bare nominal receives the theta role of patient (theme).

The derivation of this construction with Yaqui *peu* - "butcher" is given in (29):

(29) ... vP vP ...
      \   \   
      N  v   N  v
      \      \      
     maso   peu   ti   maso-peu

These cases do not allow for modifier-stranding since in the absence of DP structure there is no room for non-nominal material (adjectives, determiners, quantifiers, etc.) to appear. Once we allow for bare-N complementation we yield N-V compounding via the syntactic operation of head movement, which is triggered by the inability of the nominal to receive Case.

However, it remains the case that, in Yaqui at least, this kind of optional construction is limited to only a handful of verbs, which do not necessarily require this kind of complement, since a full DP is also an option (cf. 28a). (As pointed out in Chapter 7, it is not clear to what extent this compounding process is active in many of the Uto-Aztecan languages, despite the implicational relationship suggested by Mithun 1984 that languages with classificatory and other types of NI should also have N-V compounding).

Although this construction is an option available to a few verbs in Yaqui, the configuration in (29) is what is required to yield a direct object reading with NI of the N-V compounding type. Although the positing of such a syntactic process in Yaqui (and
other Uto-Aztecan languages with the equivalent constructions) may seem ad hoc in the
case of its non-productivity, I will show in section 8.4.2 that there is further cross-
linguistic variation with respect to what sorts of complements verbs may take. Thus does
a syntactic analysis posited for a relatively minor phenomenon in the Uto-Aztecan
languages predict language types that are in fact revealed from further cross-linguistic
inquiry, as is expected from a view of syntactic variation based on parameterization.

8.4.2. Other complements to V

The claim of the previous section is that the difference between N-V compounding and
syntactic NI lies in the complement to the verb—syntactic NI, perhaps the usual case,
involves a DP complement, whereas N-V compounding involves a bare N complement.
In this section I will briefly consider further cross-linguistic evidence that suggests that
different elements can be sister to V. Specifically, we will consider NP as the
complement to V in Niuean (Polynesian), in a construction that Massam (2001) has
termed "pseudo noun incorporation" (PNI).

Massam shows that incorporated elements in Niuean are phrasal in nature, since they
allow for modification, as in (30):³

(30) Ne kai sipi mo e ika mitaki a Sione (Massam 2001: 160 [7b])
Past eat chip Comtv Abs fish good Abs Sione
"Sione ate good fish and chips"

³ K. Hill (2003) also discusses some limited examples of incorporation of adjectival elements with
nominals in Hopi.
Under Massam’s analysis, in the PNI constructions the complement to V is an NP, which cannot receive accusative case since it does not have D or case features. This complex NP forms a predicate with the V, and “undergoes fronting to satisfy the ‘predicate first’ nature of Niuean” (p. 165), thus accounting for the VOS order of these constructions. In this analysis, it is a “maximal projection which undergoes predicate fronting”, a phenomenon that is also seen in other constructions in Niuean, such as predicate nominals and prepositional predicates (p. 165). As Massam emphasizes, this PNI construction is neither appropriately analyzed as head movement nor as lexical compounding. It is formed via the “base generation of a non-referential bare NP which, because it is not a DP, undergoes subsequent predicate fronting along with the verb of the sentence to derive a VOS word order” (p. 193).

There might be a parallel scenario in Uto-Aztecan, in the Taracahitic language Guarijío, which has properties of both a verb-final and verb-initial language. According to Barreras Aguilar (2000), sentences with indefinite direct objects have an OVS order:

(31) O V S (Barreras Aguilar 2000: 133 [52])

sesénu hente nete-ré no’ nó-a=pu te’ pá asi-sá-pa
another people make-Rem.Pst Father-Emph=PRIV arriba arrive-Adv-iter

‘The Lord created another people when he arrived on high’

I will speculate here that a similar analysis could be posited for Guarijío, with an indefinite NP complement fronting with the verb in these constructions. However, I will have to leave the details of such an analysis for future research.

The data presented in this section suggest that there is a typology of verbal complementation that includes at least DP (assuming that incorporation is allowed out of DP, at least in some languages, e.g. Hopi—see discussion in 8.2.1, and for evidence for
the referentiality of NI in Hopi see K. Hill 2003), NP, and N. I will leave a more explicit discussion of the internal structure of DP in Uto-Aztecan, and its consequences for cross-linguistic NI, to future research.

We turn now to another aspect of Uto-Aztecan NI: obligatory object polysynthesis.

8.5. Object polysynthesis and pronominal arguments

Thus far in this chapter we have found a way to formalize an identical account for three of the NI types discussed in Chapter 8—noun incorporation as syntactic head movement. Classificatory NI has been linked to syntactic NI via Late Insertion, and N-V compounding has been linked to syntactic NI via a stipulated bare N complement for V. In this section I will argue that “object polysynthesis” is similarly derived via head movement, and that it can in fact be considered an instantiation of NI with hyponomous objects, or classificatory NI.

Polysynthetic languages, under Baker (1996, 2001)’s definition, obligatorily require subject and object marking on the verb. In Chapter 9 I suggest that there is good diachronic reason to separate subject agreement from the morphological marking of the direct object, so for our purposes here we can focus simultaneously on those languages that have full-blown polysynthesis, such as the varieties of Nahuatl, and those languages that have “obligatory object polysynthesis”, such as Tohono O’odham.

The claim that I make here is that languages of these types in fact have obligatory incorporation of the direct object, which in languages like Nahuatl can involve incorporation of a nominal root or the incorporation of a pronominal element. Languages
like O’odham, on the other hand, do not allow the incorporation of nominal roots but instead require a pronominal to be incorporated. Typical examples are those such as (32) and (33):

(32) Nahuatl “object-agreement”
   a. askeman ti-‘-kwa nakatl
      never you-it-eat meat
      ‘You never eat meat’
   b. na’ ipanima ni-naka-kwa
      I always I-meat-eat
      ‘I eat meat all the time’

(33) Tohono O’odham Obligatory Object-Marking
   a. A: ūi aŋ g haiwaŋ (ŋ)-ceposid
      1sg AUX DET cow (SG.OBJ)-brand
      ‘I am/was branding a cow’
   b. A: ūi aŋ g wipsilo ha-ceposid
      1sg AUX DET calves PL.OBJ-brand.PL
      ‘I am/was branding the calves’

I claim that the syntactic difference between these two cases is that polysynthetic languages such as Nahuatl involve obligatory NI, and that when object agreement occurs with an overt object nominal (as in 32a) what we see is in fact classificatory NI. The obligatory pronominal incorporation in O’odham (33) is also classificatory NI.

Rather than just being “agreement”, as it is often construed, I claim that the elements incorporated into the verb in these constructions are pronominal classifiers. (Following the diachronic analysis developed in Chapter 9 I claim that the subject marker is agreement, triggered presumably at TP or some other functional projection higher than

---

4 The third person object agreement prefix is null for the third person singular, but is ha- in the third person plural. See 7.4.4 for the pronominal object prefixes in O’odham.
vP). In these cases this object pronominal acts as a classifier taking an overt hyponomous argument DP, and hyponomous arguments are inserted into the “trace” position of the lower “copy”, just as with the cases of classificatory NI outlined above in section 8.3.

In cases such as O’odham, where we never see overt noun incorporation, there is a further requirement that the incorporating nominal must be pronominal. I take it that this is a parameter setting that can develop from Nahuatl-style NI, the result being in effect obligatory classificatory NI. This parameter is the Pronominal Argument Parameter, described by Jelinek (2001), as applied to direct objects. These “pronominal” object elements in effect act as classifiers, in the simplest case indicating simply person and number, as in O’odham. As discussed above, some languages (e.g. Caddo) can have much more complicated classifier systems.

The difference between Jelinek’s analysis of pronominal argument languages and what is proposed here is that Jelinek’s theory of pronominal arguments precludes full DPs from appearing in argument positions. Under my analysis, direct object arguments must originate in the complement to v. After the obligatory incorporation and spell-out of the pronominal in configuration with v, hyponomous DP material may be inserted into the lower copy and from here may move elsewhere for discourse-functional purposes such as the expression of topic or focus, etc., as Jelinek’s work has so clearly shown.

The analysis sketched here leaves open a number of problems which relate to arguments that have been raised with respect to full DPs in polysynthetic and pronominal

---

5 The exception is in the case of denominal verbs, where incorporation is forced due to the “defective” nature of the verbal head, which has a morpho-phonological requirement that it be an affix.
argument languages, which have hitherto been analyzed as having DPs in A’-positions (Baker 1996; Jelinek 1984, 2001). Since my focus is entirely on the relation of NI to hyponomous direct objects, I must leave some of these questions unanswered. I leave open the possibility that the analysis that I have presented is entirely diachronic in nature, and that if Jelinek is correct then the insertion of DPs into adjunct positions has come about only after the grammaticalization of obligatory NI in these languages, where incorporated direct object pronominals have been reanalyzed as agreement morphemes.

Recent research suggests that there may actually be a continuum of incorporated pronominals and object agreement. Lichtenberk (1997) argues that object-marking is polyfunctional in To’aba’ita (Austronesian), in that it serves as agreement whenever there is an overt pronoun or NP object, but that the object-marker itself constitutes the argument when there is no overt object. Baker (2003) argues that object agreement on the verb only appears if object NPs are in dislocated (A’-) positions in the Bantu languages Chichewa and Kinande.

If it turns out that “object agreement” must be separated from actual head-moving noun incorporation, one possible line of argumentation for the diachronic development of this arrangement could be made along the lines of Roberts and Roussou (2003), who present a Minimalist Program approach to grammaticalization wherein movement (e.g. incorporation) is dispreferred by the language-learner, under the assumption that parameter-setting is driven by simplification of syntactic representations.

I leave for future research whether or not the ideas laid out here can be applied to other languages that have figured prominently in debates on the configuration of these
types of languages: Warlpiri, Mohawk, Lummi, Navajo, etc. The diachronic development of pronominal argument and polysynthetic languages in Uto-Aztecan will be the focus of Chapter 9.

To conclude this section, I have claimed that polysynthetic languages such as Nahuatl display obligatory NI. Pronominal argument languages in Uto-Aztecan, on the other hand, have obligatory NI with a further morphological requirement to spell out a hypernomous item (i.e. classifier) in the N-V complex.

8.6. Conflation and non-object noun incorporation in Uto-Aztecan

In section 8.4 we focused exclusively on those cases of NI that involve incorporation from direct object position, where we see patient theta-roles associated with the incorporated nouns. My argument in that section was that the theta-role of theme is assigned in a specific syntactic position, complement to V, which in and of itself is a relatively uncontroversial position. The novel claim is that even cases of apparent compounding NI, with detransitivization of the verb, are derived via head movement from this position.

In this section we will consider non-object NI. These include what I will term “manner conflation”, which derives instrumental prefixes and other adverbial modifiers in complex verb-formation, as well as cases of apparent “subject incorporation”. Each of these involve the merging of a root with a verbal head, but the question is this: where does this merger take place?
8.6.1. Instrumental prefixes and manner conflation

Examples of instrumental prefixes in Uto-Aztecan were given above in 7.3.4.1.

Instrumental prefixes typically perform one of three functions: identifying the means (instrument) with which the action denoted by the verb is performed, the manner in which the action denoted by the verb is performed, and occasionally the patient (i.e. direct object) of the action of the verb. In some cases IPs can also be used as directionals (Thomes 2000: 256). Some usages of instrumental prefixes may in fact be ambiguous between these functions.

In section 7.3.4.1 above I suggested that "instrumental prefixes" functioning as a direct object of a verb originate in direct object position, and are derived through incorporation, just as is the case in standard noun incorporation. The lack of ability to modify the instrumental prefix would indicate bare N complementation, just as argued above in section 8.4.1. for intransitivizing noun incorporation in Yaqui.

The discussion of instrumental prefixes (IPs) to follow in the remainder of this section is based largely on the recent work of Thomes on IPs in Northern Paiute (Thomes 2000 and 2003). Thomes (2000) examines Northern Paiute instrumental prefixes from a typological perspective, comparing IPs in Northern Paiute to those that are found in other North and South American languages, including Lakhota (Siouan), Klamath (Penutian), Haida, and Bora (Huitotoan). He found that IPs have very similar forms and functions across these diverse languages, and I assume that the properties that he elucidates in detail for Northern Paiute apply to instrumental prefixes much more generally, as well as to the instrumental prefixes in closely related Uto-Aztecan languages more specifically.
As noted above, Langacker (1977a: 133) suggests that instrumental prefixes are largely limited to Numic and Tepiman within Uto-Aztecan. DeLancey (1996) locates Numic in the so-called “bipartite stem belt”, an areal pattern of stem formation in which (primarily) verbs obligatorily require two morphemes, often including instrumental prefixes. Numic differs from most the central members of this areal pattern, however, in that

the core languages are characterized by the fact that a majority of the verb stems of the language are bipartite, consisting of a L[exical] P[refix] and a second element. The second element is not always a L[ocative]-D[irective] S[tem], but typically most or all of the potential second elements in a language are bound morphemes which cannot occur without a LP. (DeLancey 1996)

The very limited number of “instrumental verbs” in the Numic languages make them very different from the “core” languages of the bipartite stem belt in this respect. However, DeLancey concludes with the idea that Numic must have been a part of this areal spread at an early period, contrary to some recent claims about Uto-Aztecan prehistory (e.g. J. Hill 2001, 2004; Dakin 2003, 2004), and, even more controversially, points to a suggestion by Jacobsen (1966) that Numic may have been the source for this areal pattern. However, the Tepiman languages cannot plausibly be regarded as members of this areal spread, and it is possible that IPs in these languages are retentions from Proto-Uto-Aztecan (as argued by Sherzer and Foley 1971).
8.6.1.1. Instrumental prefixes in Northern Paiute

Discussions of IPs often relate these morphological pieces to noun incorporation, since NI is often taken to be the etymological source for IPs. In his discussion of “Incorporation” in Uto-Aztecan, for instance, Langacker (1977a) states that “despite their limited distribution, instrumental prefixes are the best known of the incorporated elements in UA and have received the most attention” (p. 133). However, in considering the characteristics that distinguish IPs from the true noun incorporation constructions involving head movement which have been the focus of much of the present chapter, we come to the conclusion that IPs serve adverbial rather than strictly nominal functions. There are several lines of argumentation that lead to this conclusion, and to the differing syntactic analysis which follows from it.

Although perhaps not the most convincing, the first is historical. Although many IPs (in Northern Paiute as well as cross-linguistically) have their etymological roots in nominals, IPs with etymological roots in verbs are not unattested, and Thornes (2000: 265-6) suggests that IPs “functionally correlate” to verb serialization constructions, and that some synchronic IPs may have been historically derived from such constructions. Thus, “instrumental” prefixes can have sources that are not nominal, a point that has been overlooked in much previous work.

In reference to the synchronic adverbial nature of IPs, Thornes points out that “the use of the term ‘instrumental’ can in fact be misleading, since it is more often some feature of the manner in which the action is carried out which is indicated, not the nature of the instrument” (p. 357). For example, consider the pair of verbs in (34):
Regarding the difference in meaning of these examples Thornes states the following:

Here the contrast is not in the shape or type of instrument used—one could use a wooden spoon for either task—but in the nature of the action. In [34a], the end of the instrument is used, while in [34b], the preparation of the gravy requires the instrument to be held at an angle closer to the horizontal in order to blend the thickener and prevent scorching. (p. 357)

Since the IPs used with these verbs do not place a classificatory restriction on external “instrumental phrases” used with these constructions, Thornes considers these elements to be adverbial. Thus, our syntactic account must provide a way to morphologically add adverbial material to the verb, rather than actually “incorporate” an instrumental nominal.

In addition to a difference in function vis-à-vis nominals, however, these elements also have some other important properties that should be pointed out. First, these elements are non-referential—external instrumental phrases may be included by a speaker in a given sentence with an IP in order to “be explicit about the particular instrument involved in the action” (p. 357), but the IPs themselves are not referential.

Some IPs have developed into a morphological causative, where the addition of the IP leads to a transitive verb stem, as with the prefix ma- ‘hand’ (35b):

(34a) 
\[
\begin{align*}
  {tsi-kwidu} & \; {\text{IP/sharp-stir}} \\
  & \; \left\{ \text{stir (e.g. soup)} \right\}
\end{align*}
\]

(34b) 
\[
\begin{align*}
  {wi-gwidu} & \; {\text{IP/long-stir}} \\
  & \; \left\{ \text{stir (e.g. gravy)} \right\}
\end{align*}
\]

(Thornes 2003: 357 [7])

(35a) 
\[
\begin{align*}
  {u-su} & \; {s\tilde{i}\tilde{a}-kwi} & \; {\text{3-NOM be.afraid-CONT.SG}} \\
  & \; {\text{‘S/he is afraid’}}
\end{align*}
\]

(35b) 
\[
\begin{align*}
  {u=} & \; {ma-z\tilde{i}\tilde{a}-u} & \; {\text{3=CAUS-be.afraid-PNC}} \\
  & \; {\text{‘(S/he) scared her/him’}}
\end{align*}
\]

(Thornes 2003: 368 [22])
These data support the verb-serialization analysis—in these cases the “IP” is realizing the manner of causation, i.e. is spelling out a ‘v’.

However, not all instances of this prefix are used as a causative. For example, Thomes states that “the following pair indicates that with the verb root {yui} ‘warm, be warm’ it does not have this effect” (2003: 361), as in (36)a, which requires the applicative to derive a causative reading (36)b:

(36)a. ma-yui-kati  
b. i=mai yui-ki-kati  
   IP/hand-warm-sit  
   ‘sit hand-warming’  
   (Thomes 2003: 361 [11])

(36)a looks like actual noun incorporation, where the nominal root indicates the patient argument of the verb. Thomes notes the existence of similar examples where the “IP” “exhibits a pattern more akin to noun-incorporation, resulting in an intransitive stem” (p. 362), such as (37)e, which contrasts with “the more common means-and-manner pattern whose result is transitive”, as in (37)a-(37)d:

(37) -kyoda ‘lift’  
a. ma-kyoda  ‘pick up’  
b. ta-kyoda  ‘lift with toes/feet’  
c. ts-a-kyoda  ‘pull someone up’  
d. tsi-kyoda  ‘lift with the tip of something’  
e. tso-kyoda  ‘raise head up’  
(Thornes 2003: 362 [13])

However, there are also cases of true syntactic ambiguity between IP and NI constructions. According to Thomes, speakers vary as to whether or not they use IP constructions productively—“in Northern Paiute, younger speakers seem to have a stronger tendency to treat IP-Root combinations as lexical items, and therefore simply do not make use of their creative potential in word formation, a skill older, more fluent speakers retain” (p. 364). In addition, “speakers also readily create a context for unusual
or 'invented' IP-Root combinations or may interpret the same stem differently, based on context" (p. 365). According to the theoretical proposals made in this chapter, in those cases where the IP is interpreted as a direct object the hearer must reconstruct NI of the bare nominal type.

This potential for ambiguity seems to belie Thornes' previous (2000) characterization of the IPs of Northern Paiute, which he lists as the following (IPs with a fortis/geminating feature are indicated with '[-]'):

(38) Instrumental Prefixes in Northern Paiute

<table>
<thead>
<tr>
<th>IP</th>
<th>Etymological Source</th>
<th>Gloss</th>
<th>Characterization</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. wi[-]</td>
<td>UA *wipaa 'whip'</td>
<td>'long'</td>
<td>along or against the length of a long object; radial motion; natural forces</td>
</tr>
<tr>
<td>b. tsi[-]</td>
<td>PN *tsi-a 'rose' ?</td>
<td>'sharp'</td>
<td>with a sharp object; with the end of a long object</td>
</tr>
<tr>
<td>c. ki'</td>
<td>UA *ki?i 'bite'</td>
<td>'bite'</td>
<td>with the teeth; by biting; edge or periphery</td>
</tr>
<tr>
<td>d. ma[-]</td>
<td>PN *mai, UA *mo?o</td>
<td>'hand'</td>
<td>with the hand (usually open); causative; with respect to hand or arm</td>
</tr>
<tr>
<td>e. to[-]</td>
<td>???</td>
<td>'fist'</td>
<td>with the fist, axial motion; out from the body</td>
</tr>
<tr>
<td>f. tsa'</td>
<td>PN *tsa?i 'grasp'</td>
<td>'grasp'</td>
<td>with the fingers; grasping; toward the body</td>
</tr>
<tr>
<td>g. ta[-]</td>
<td>PN *ta?ga, UA tannah</td>
<td>'foot'</td>
<td>with the feet/leg (verbal); stepping; w.r.t. feet/leg</td>
</tr>
<tr>
<td>h. ku[-]</td>
<td>UA *kuh 'fire'</td>
<td>'heat'</td>
<td>with heat; w.r.t. fire, smoke</td>
</tr>
<tr>
<td>i. ni[-]</td>
<td>?? ni-a 'call'</td>
<td>'speak'</td>
<td>with speech; by talking</td>
</tr>
<tr>
<td>j. mu[-]</td>
<td>PN, UA *mupi 'nose'</td>
<td>'nose'</td>
<td>having a point; w.r.t. nose/beak, lips/mouth</td>
</tr>
<tr>
<td>k. pu-/pui</td>
<td>[??]</td>
<td>'eyes'</td>
<td>with the eyes, by seeing</td>
</tr>
<tr>
<td>l. si[-]</td>
<td>UA *sUp 'cold'</td>
<td>'cold'</td>
<td>cold; fear; shaking</td>
</tr>
</tbody>
</table>

(Thorones 2000: 271-2 [Appendix])
<table>
<thead>
<tr>
<th>Prefix</th>
<th>Root</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>m. pa[']-</td>
<td>UA *paa</td>
<td>'water' pertaining to water, moisture; N-I with /paa/</td>
</tr>
<tr>
<td>n. tso[']-</td>
<td>UA *tsohji</td>
<td>'head' head/shoulders</td>
</tr>
<tr>
<td>o. pi[']-</td>
<td>UA *pih</td>
<td>'butt' buttocks/back/behind/base; usu. as a location</td>
</tr>
<tr>
<td>p. ti [']-, to-, ta-</td>
<td>UA *tUn</td>
<td>'rock' with a rock; forceful activity, like 'fist'</td>
</tr>
<tr>
<td>q. ku/ko[']-</td>
<td>PN *koba-i</td>
<td>'face' face, neck, above, in front</td>
</tr>
<tr>
<td>r. hu-</td>
<td></td>
<td>'back' back; outside; behind; typically location</td>
</tr>
<tr>
<td>s. su[']-</td>
<td>UA *sunna</td>
<td>'mind' mental or emotional activity</td>
</tr>
<tr>
<td>t. ta-</td>
<td>PN *taba</td>
<td>'sun' sun/day/light; appear</td>
</tr>
<tr>
<td>u. no-</td>
<td>PN *no-</td>
<td>'load' egg/house/round; move/carry</td>
</tr>
</tbody>
</table>

In his discussion Thomes (2000) distinguishes between four large groupings of these prefixes, which I have divided into (38)-(38''). This arrangement is intended to reflect “subclasses based on transitivity considerations—whether or not they have a transitiivizing effect on the outcome of the derived stem—as well as by their general productivity”. The IPs in [38] “are by far the most productive of the set”, and these have the “transitiivizing effect” mentioned above. The IPs in [38'] “are more restricted in their productivity, likely for semantic reasons, but typically occur as part of the derivation of transitive verb stems”. The IPs in [38''] “pattern like incorporated nouns... and typically derive intransitive stems”. And IPs in [38'''] “are very restricted, occurring in only a few stems, and as such appear to be lexicalizations” (p. 263). Perhaps it is the case that ambiguities in the use of these IPs are limited to the more fluent, older speakers whom Thomes identifies as more productive in their use of the IPs.
In the next section I present a syntactic analysis of adverbial IP constructions, invoking Hale and Keyser (2002)'s notion of conflation.

8.6.1.2. A syntactic analysis of instrumental prefix constructions

Recall from section 8.6 that conflation is a property of Merge, whereby some Root is merged with a morphological head, in this case v. This Merge happens under “strict complementation.”

The analysis that I propose for these verbs is based on Harley (in press)'s recent proposal for denominal verbs in English, what she terms “manner incorporation.” Because I follow Baker (1988)'s terminology in regarding incorporation as head movement, I will employ Hale and Keyser's term conflation in the context of Harley’s theoretical proposal.

The essential question at hand here has to do with the following query, in Harley (in press)'s phraseology—how do verbs “get their names?” Or, in other words, how do verbs get their phonological content? One source, the focus of most of this chapter, is noun incorporation, which I have argued is instantiated via head movement in both verbal constructions involving an overt lexical root with NI as well as in denominal verb-formation.

In traditional grammar, we would posit a structure like (39) for such verb phrases as dance a jig:
Under the traditional account, there is a verb *dance*, which is independent of the noun *dance*. Thus, the Root *dance* has two categorial entries: *dance*$_N$ and *dance*$_V$, which are, in a sense, only coincidentally related; they are certainly not derivationally related. What a Hale and Keyser-style theory suggests is a more radical view, but one consistent with recent conceptualizations of word-building within Distributed Morphology. This is that the bifurcation of what is clearly the same Root, *dance*, into two categorial sub-headings is spurious. Rather, what we have is the same Root, *dance*, which can be licensed in multiple syntactic “slots”, which yield the results of nominal or verbal category. From this perspective nominal or verbal category status is an epiphenomenon of location within syntactic structure (Pesetsky 1995, Harley and Noyer 1998). In the case of English *dance*, in the complement position of a causative light verb (v), i.e. in nominal direct object position, $\sqrt{dance}$ would appear as a noun. When incorporated into little v, dance is by definition a verb.

With hyponomous arguments, as in *dance a jig*, I have argued that the verb gets its features via the insertion of a root into the head of the movement chain created by syntactic noun incorporation, and that the lower copy may be spelled out with different morphological material as long as equivalent features are involved. This derives the hyponomous relationship between an incorporated head and its coindexed complement. This analysis only applies to those cases of verb-formation that involve direct object
arguments, including "instrumental prefixes" serving the function of the theme of the verb. There is another possible source of verb-formation, however, and this involves the "conflation" of means and manner adverbial elements into the v position.

Harley (in press) argues that for English instrumental denominal verbs (hammer, skateboard, etc.), in which the root of the verb represents an instrumental adjunct, the roots have no obvious source in sentential argument structure. These verbs cannot be the result of a root incorporating into v, because there is no place within the argument structure of the verb from which these nominals could incorporate, adjuncts being A'-positions. Rather, they must receive their phonological material from manner conflation. In this process, "a v may be named by a Root describing the Manner in which it is accomplished" (p. 26). Further, "assuming that all adjuncts, including Instrumental ones, are a species of Manner, these denominal verbs represent an occurrence of Manner [Conflation] applying to an 1-syntactic structure that would normally give rise to a verb of contact, involving a complement headed by an Event-denoting Root" (p. 26).

Harley illustrates the syntax of manner conflation with the following kind of notation:

(40) ... vP
     △ Sue v' vP
     (hit) the metal

This notation is used to suggest that the phonological properties of the verb are derived from the Merging of means/manner root directly into v, via Late Insertion, without
having to derive the means or manner element directly from the argument structure. As Harley points out, this kind of verbalization of instrumentals is completely productive in English.

Extending this notion to Northern Paiute IP constructions, we note that some verbs can take IPs optionally, whereas other verbs, the so-called “instrumental verbs”, are bound roots that require IPs, such as the following: -ka? a ‘to cut, sever, split’ and -bosa ‘to pierce’ (Thomes 2003: 367). According to Thomes, “this class of bound roots in Northern Paiute appears most often to be of the change-of-state type” (2003: 367).

The analysis that I propose for manner IPs is the following. Some verbs in Northern Paiute (and other languages with similar constructions) subcategorize for IPs, either obligatorily (as with the closed class of “instrumental verbs”) or optionally. Whereas English allows for practically any nominal that can serve an instrumental function to do so as a verb, Northern Paiute has a small, closed class of elements that can be added to an overt v element, and some v’s which are “defective” in some sense and that require this additional support. The syntactic addition of this extra morphological material is carried out by the same mechanism that creates compounds—Merge. If the IP is not added to the verbs that obligatorily take IPs then the relevant manner features will not be “checked” and the derivation of the verb in question will “crash.” This Merging process is akin to compounding.

The “compounding” analysis is further supported by the fact that the roots that constitute the class of “instrumental prefixes” on verbs can also be used in combination with nouns. As Thomes puts it: “IP’s occur as formatives for terms relating to body parts
and other stems that otherwise have the morpho-syntactic properties of nouns in Northern Paiute. Under these circumstances, they have a classifying effect” (p. 112). The contrast between the two sets of words following with the prefixes ma- ‘of or pertaining to the hand’ and ta- ‘of or pertaining to the foot’ are typical:

(41)a. ma-ñwitsogo a’. ta-ñwitsogo
   ‘wrist’ ‘ankle’

b. ma-sihanî b’. ta-sihanî
   ‘fingers’ ‘toes’

c. ma-togo c’. ta-togo
   ‘thumb’ ‘big toe’

I would argue that these kinds of elements are derived via compounding, just as is the case with IPs conflating with verbs. The only difference between this process and compounding as more generally conceived is the obligatorily bound nature of these prefixal elements, which only occasionally have free root equivalents in any given language. I follow the general track of received thought in Uto-Aztecan linguistics and grammaticalization theory that claims that such bound elements have their historical origin in free roots (e.g. Mithun 1991), and over time they have been reduced to their current affixal state, thus being limited to use with one syntactic process—conflation (i.e. compounding), implemented by Merge.

8.6.2. Subject incorporation

The issue of the existence of subject incorporation has vexed linguists since at least Kroeber (1909), who had argued, unsuccessfully, that noun incorporation could not exist
because of the asymmetry between subjects and objects in incorporation constructions. In Kroeber (1909)'s view, "a theoretical objection to objective noun-incorporation is the fact that no one has affirmed incorporation of the subject. And yet there is not more reason why the object should be fused with the verb than the subject, and if objective incorporation is found in many languages, subjective incorporation should at least occur sometimes" (p. 573).

Baker (1988) and all subsequent syntactic approaches to the issue of noun incorporation have precluded the incorporation of *agentive* subjects in principle, via some theoretical mechanism such as the Head Movement Constraint (Travis 1984). In the head-movement view of incorporation, agentive nominals cannot incorporate because this would require downward incorporation from a specifier position, in violation of constraints on movement.

Still, however, this issue remains an empirical one, and various scholars have presented evidence that they claim constitutes subject incorporation. Such evidence has come from a variety of languages, including Koyukon Athabaskan (Axelrod 1990) and Lakhota (de Reuse 1994). In Uto-Aztecan, subject incorporation has been reported for Nahuatl (Merlan 1976, Tuggy 1986, K. Hill 2003) and Hopi (K. Hill 2003). It is necessary for theoreticians to account for these cases in their theories.

I divide "subject incorporation" into two general sub-groups. The first can be analyzed following the unaccusative hypothesis of Perlmutter (1978), where the incorporating "subject" is in fact a logical object, and it incorporates from the underlying object position of sister-to-V (8.6.2.1). The second case of "subject incorporation" leads
to possible “agentive” interpretation, but, I claim, is the result of simple N-V compounding, where the nominal does not derive from movement from underlying object position (8.6.2.2).

8.6.2.1. Unaccusative incorporation

Since unaccusative constructions involves movement of the incorporating element from an underlying object position, these constructions are entirely consistent with the syntactic movement account of incorporation (Baker 1988: 88-92). These can be derived as in the following example from Tetelcingo Nahuatl:

(42) Unaccusative Subject Incorporation in Tetelcingo Nahuatl

a. toonal-kiisa
   sun-emerge
   ‘the sun comes out’

b. 

\[ \begin{array}{c}
\text{toonal-kiisa} \\
\text{sun-emerge}
\end{array} \]

\[ \Rightarrow \]

\[ \begin{array}{c}
\text{toonal-kiisa} \\
\text{sun-emerge}
\end{array} \]

Other examples of unaccusative incorporation constructions in Nahuatl include the following:

(43)a. tla-α:weci-α-∅
    INDEF-water-fall-PRES-SG
    ‘It is raining’ (lit. ‘water is falling’)

b. tla-se-weci-α-∅
    INDEF-snow-fall-PRES-SG
    ‘It is snowing’ (lit. ‘snow is falling’)
(Merlan 1976 calls the examples in [43] “lexical incorporation”, since these “compounds of noun stem and verb stem express lexical meanings which cannot be rendered otherwise” p. 183, i.e. these meanings cannot be expressed with these roots without noun incorporation). Further examples are in (44) and (45):

(44) Ni̧mačiliya, nama aš čačaonka (Merlan 1976: 185 [17B])
1SG-it-regret, today NEG pitaya-BE (THERE)
‘I’m sorry, there is no pitaya today’

(45) tlaške nokʷenta tla wewemĩki (Merlan 1976: 189 [27B])
what 1SG-affair if (3SG-old) man-died
‘What do I care if (her old) man died!’

A similar unaccusative analysis could probably also be posited for verbs of motion, as with the Hopi and Nahuatl proper name examples from K. Hill (2003) (e.g. Hopi Hon-wari, bear-run, ‘The bear ran’), presented in 7.3.4.2. Such an account would not be surprising from a semantic point of view since verbs of motion can plausibly be construed as either agent- or patient-like motion along a path. The possible construals of agent-ness or patient-ness may be language-specific.

However, Baker (1988) gives the following ungrammatical examples as support for his incorporation account, since incorporation with motion verbs in these languages is ungrammatical. The ungrammatical (b) examples below seem to be equivalent to the grammatical Uto-Aztecan examples cited by K. Hill (2003) and Hale and Keyser (2002):

---

6 The reference to ‘her old man’ in this sentence had been established by a previous statement in the discourse to the effect of Do you know that her old man (husband) died yesterday? (Merlan 1976: 189 [27A]).
(46) Southern Tiwa (Kiowa-Tanoan)

a. Khwien-ide ə-teurawe-we  
   dog-SUF A-run-PRES  
   ‘The dog is running’

b. *ə-khwien-teurawe-we  
   A-dog-run-PRES  
   ‘The dog is running’

(47) Onondaga (Iroquoian)

a. H-ate-ʔse:-ʔ neʔo-tsiʔt-aʔ  
   3MS-REFL-drag-ASP the-PRE-louse-SUF  
   ‘The louse crawls’

b. *H-ate-tsiʔkiʔt-ʔse:-ʔ  
   3MS-REFL-louse-drag-ASP  
   ‘The louse crawls’

Baker notes that “[46b] may be ruled out independently in Southern Tiwa by an
animateness restriction, which says that animate subjects never incorporate (although
animate objects do). . . [t]here is much overlap between animateness and agentivity in the
subject position, but some residue of this animacy condition may have to be stipulated”
(1988: 453). If there are differences in semantically-based restrictions among these
language which allow for some to allow incorporation with verbs of motion while other
do not, they have not hitherto been discovered (as far as I know).

I leave open the possibility that incorporation of subjects with motion verbs may also
provide an instantiation of N-V compounding with agentive interpretation, as discussed
in the following section.
8.6.2.2. N-V compounding with “agentive” interpretation

Most of the “subject-incorporating” structures that have been discussed thus far involve verbs of motion, which might conceivably lead to an unaccusative analysis of these constructions, where the incorporating “subject” derives from an underlying object position. However, there are attestations of subject incorporation which possibly have a more agentive interpretation for the incorporated nominal, including the Nahuatl personal name from which we have derived the Anglicized “Montezuma”, in (47):

(48) Mo-tēuc-zōmah
    REFL-lord-frown
    ‘The lord frowned in anger.’

To expand this discussion beyond issues in Uto-Aztecan, I should point out that similar data have been discussed in unrelated languages indigenous to North America. De Reuse (1994) gives a variety of forms from Lakhota (Siouan) which purport to show agentive subject incorporation. Many of these involve motion, as well as “weather and atmospheric phenomena”. As de Reuse points out, “admittedly, the subjects in [many of the examples] are not stereotypical agents, being nonhuman, inanimate, and incapable of volition. However, they have a force of their own” (p. 220). Some of the forms are plausibly unaccusative in nature, having such glosses as ‘for the ice to be gone’, ‘for a cloud to burst’, ‘for the wind to whistle’, ‘to rain’, ‘for it to snow’, and so forth (pp. 220-1). However, de Reuse does give some forms which do appear to be agentive, even if only metaphorically so:
(49) \( \text{ph\'eta ay'\'i\'iski} \) (de Reuse 1994: 220 [92])

‘for a fire to rejoice on someone, i.e. when sparks leap in one’s direction, a sign of good luck’

(50) \( \text{Wakiy\'a ho\'thus pi} \) (de Reuse 1994: 221 [102])

‘for it to be thundering’ (lit. ‘for Thunder Beings to call’)

(51) \( \text{Wakiy\'a tuw\'a pi} \) (de Reuse 1994: 221 [104])

‘to flash lightning’ (lit. ‘for Thunder Beings to open their eyes’)

In Koyukon Athabaskan, we see not only cases of apparent agentive subject incorporation, but also, apparently, agentive subject idioms (cf. Marantz 1997).

According to Axelrod (1990),

nouns which can incorporate into a subject role typically filled by an agent noun are those inanimates which Chafe (1970: 109) has described as [+potent], that is, those which “have a force of their own which enables them to ‘perform’ certain actions.” A noun of this type “has, or is conceived to have, its own internal power.” (Axelrod 1990: 184)

Axelrod gives the following examples of nouns that can incorporate in this manner:

(52) \( \text{haatsehyedeeLtaanh} \) (Axelrod 1990: 185 [11a])

\[ hU + aa \# tseh \# ye + dee + L + taanh \]
area + PP # tears # 03sg + M/A + CL + sg/du.anim.lies
‘s/he went away crying’; lit. ‘crying [tears] carried her/him off’

(53) \( \text{kk'o'eLts'eyhyeedzoyh} \) (Axelrod 1990: 185 [11b])

\[ kk'o \# eLts'eyh \# ye + ee + \O + dzoyh \]
around # wind # 03sg + M/A + CL + multiple objs.
‘the wind is moving them around’

(54) \( \text{nonobeLyeeten'oyh} \) (Axelrod 1990: 185 [11c])

\[ no \# no \# beL \# ye + ee + te + ne + \O + 'oyh \]
down # ITER # sleep # 03sg + M/A + CL + compact obj.
‘s/he is nodding from drowsiness’; lit. ‘sleep is causing it (her/his face) to go down repeatedly’
Furthermore, although it is usually the case that animate nouns cannot incorporate, “inanimate nouns cannot function as agent with verbs that normally take a volitional agent unless they are incorporated” (p. 186, emphasis in original), as in (56)b:

(56)a. John hehetaatltaanh

\[ J \quad \text{hebe} \quad + \text{te} + \text{le} \quad + \text{L} \quad + \text{taanh} \]
\[ J \quad \text{O3pl} \quad + \text{M/A} \quad + \text{CL} \quad + \text{sg./du.anim.lies} \]
‘John carried them off’

b. tohebtaatltaanh

\[ \text{to} \quad \# \quad \text{hebe} \quad + \text{te} + \text{le} \quad + \text{L} \quad + \text{taanh} \]
\[ \text{water} \quad \# \quad \text{O3pl} \quad + \text{M/A} \quad + \text{CL} \quad + \text{sg./du.anim.lies} \]
‘they floated away’; lit. ‘water carried them off’

c. *too hebtaatltaanh

Axelrod summarizes these facts as follows:

Incorporation, then, allows a noun which ranks low in the scale of naturalness as an agent of a transitive verb to function as the non-agentive subject of that verb. That is, where the semantic characteristics of a noun make it otherwise ineligible for agent status with a particular verb, it can be incorporated and function as a ‘nonintentional’ causer. The incorporated position, then, denotes [-control] for the incorporated subject. (p. 187)

Experiencer” (ObjExp) verbs (e.g. frighten, amuse, etc.) involves a Causer theta role assigned to a position lower than v, by which arguments of Causer type are “internal” to VP. Chtareva (2003) applies this analysis to subject idioms in Russian, wherein Causer subjects are able to form idioms with their verbs (cf. Marantz 1997).

It seems that a similar analysis is available for the NI facts of Koyukon Athabaskan, where the incorporate is a causer, but it cannot be agentive. Perhaps we can associate Axelrod’s feature [+control] to the position of specifier of v, which in Koyukon can only be occupied by nouns with the feature [+animate]. In this way we can account for these examples of subject incorporation while simultaneously maintaining Marantz (1997)’s generalization regarding agentive subject idioms.

Returning to the Uto-Aztecan data, the solution to the puzzle of agentive subject incorporation that I will propose here is that these instances cannot be derived via head movement, as Baker’s theory suggests. However, as has been repeatedly emphasized in this work, there is another way to generate an N-V “incorporation” construction, and this involves conflation. In the terms of traditional grammar, such constructions are created by simple noun-verb compounding. Thus, in these examples I hypothesize that the construction is simply the result of the Merging of a nominal root into the verbal position, rather than derivation by head movement. This compounding leads to an idiomatic interpretation wherein the incorporated nominal is taken to be the subject of its verb. As Mithun (1984) has amply illustrated with examples like English alligator shoes, such compounding does not force a particular interpretation, so idiomatization can lead to
agentive readings in some cases, and direct object, instrumental, or other manner readings in others.

This should not be possible with "syntactic NI", where the incorporating N is derived from underlying object position leaving a trace or allowing for the insertion of a hyponomous argument. Thus, agentive constructions with N-V compounding should not be able to have "stranded modifiers" for the "incorporated" agentive subject. This seems to be true for all of the Uto-Aztecan examples that I have examined, but this analysis predicts that this should be true for such constructions cross-linguistically.

8.7. Conclusion

In this chapter I have given a unified account of the four Uto-Aztecan NI types discussed in Chapter 7. Incorporation of nominals with a theme theta-role involves head-movement of the object nominal into the verbal position from the position of complement to V. "Intransitivizing" NI involves a bare N complement, and classificatory NI involves the spelling-out of the lower copy with lexical material instantiating the same features that are spelled-out by the incorporated noun. Non-object NI, as with instrumental prefixes and "subject incorporation", involve Merge alone (i.e. compounding).
CHAPTER 9

ON THE GRADUAL DEVELOPMENT OF POLYSYNTHESIS IN NAHUATL*

9.1. Introduction

One of the widest-known and best-elaborated proposals for a macroparameter within the Principles and Parameters approach to syntactic variation is Baker (1996, 2001)'s “Polysynthesis Parameter”, which is very closely related to Jelinek (2001)'s “Pronominal Argument Parameter”. In both of these proposals certain languages require some kind of marking for both subject and object arguments on the main verb of a sentence. The macroparametric approach has been criticized along several lines, not the least of which is that it seems to presuppose an abrupt acquisition of the parameter in question, i.e. a “catastrophic” change in grammar brought on by a new generation of speakers (acquirers) of a given language. This approach to parametric change is perceived to be inconsistent with other, more clearly-established aspects of historical syntax, e.g. the gradual grammaticalization of case-markers from adpositions (Trask 2002).

Although Baker’s work is notable for the breadth of its coverage of typologically distinct languages, one area that he has not focused on is in giving detailed attention to specific historical situations. That is, Baker has not spent much time looking at languages that are genetically-related to the polysynthetic languages he has studied but which are

* An earlier version of this chapter was presented at the 8th Diachronic Generative Syntax Conference (DIGS VIII) at Yale University on June 26, 2004. I thank that audience for valuable questions and comments.
not themselves polysynthetic. This is important because the parametric view posits certain factors that correlate with the parameter in question, although in many cases it has not been established whether those correlations are caused by or lead to the development of the parameter in question.

In this chapter I consider one well-known polysynthetic Uto-Aztecan language, Nahuatl, in diachronic and comparative perspective. Although Nahuatl is indeed polysynthetic, it is unusual in this regard with respect to other Uto-Aztecan languages. Based on considerations from comparative Uto-Aztecan morphosyntax I will argue in favor of the gradual (stepwise) rather than catastrophic development of polysynthesis in Nahuatl. The key factor in this gradual development was noun incorporation (NI), as discussed in earlier chapters of this dissertation.

This chapter is structured as follows. In section 9.2, I discuss the notion of polysynthesis as it has been applied from several differing perspectives, and elaborate on Baker’s macroparametric proposal as well as a recent critique of it based on empirical evidence from Nahuatl (MacSwan 1998). In section 9.3, I make a distinction between subject and object pronominal marking, showing from languages in Uto-Aztecan and elsewhere that the two should be distinguished. In section 9.4, I discuss the syntactic structure of Nahuatl relevant to Baker’s definition of polysynthesis, and in section 9.5 I put these facts into the larger Uto-Aztecan context, showing that many of the aspects of Nahuatl have “precursors” across the family. Section 9.6 concludes.

---

1 It should be pointed out that this is not possible with his work on Mohawk, since all of the Iroquoian languages are polysynthetic.
9.2. Perspectives on Polysynthesis

The notion of "polysynthesis" has been used in a variety of ways by different linguists. According to Fortescue (1994), the use of the term "polysynthetic" as a linguistic type like "isolating", "agglutinative", and "fusional" has been attributed to von Humboldt's paper from 1836, but von Humboldt's German term *einverleibend* "is usually translated as 'incorporating.' The American linguist W. D. Whitney appears to be the first to have used the word 'polysynthesis' as such" (p. 2600).

Boas (1911) characterized polysynthetic languages as those languages wherein "a large number of distinct ideas are amalgamated by grammatical processes and form a single word, without any morphological distinction between the formal elements in the sentence and the contents of the sentence" (cited by Drossard 1997: 251). In describing these languages, Boas pointed out that transitive verbs in some polysynthetic languages incorporate the object of the verb, but also that this is not a universal feature (cf. Haida).

In his (1921) discussion of language typology, Sapir defines polysynthesis as follows:

A polysynthetic language, as its name implies, is more than ordinarily synthetic. The elaboration of the word is extreme. Concepts which we should never dream of treating in a subordinate fashion are symbolized by derivational affixes or "symbolic" changes in the radical element, while the more abstract notions, including the syntactic relations, may also be conveyed by the word. A polysynthetic language illustrates no principles that are not already exemplified in the more familiar synthetic languages. It is related to them very much as a synthetic language is related to our own analytic English. (p. 128)
Sapir cautions that the terms *analytic, synthetic,* and *polysynthetic* "are purely quantitative" and relative (p. 128), and, ultimately, "too merely quantitative" (p. 136) for the purpose of determining the types of "genius" (i.e. underlying structural plans) available to different languages.

In more recent work, Drossard (1997) offers a useful typology of polysynthetic languages that provides a useful point of departure for our discussion of the development of polysynthesis in Nahuatl. In many languages regarded as "polysynthetic", a large number of adverbial elements may accrue to a verb stem, but person-marking must occur by means of free pronouns. One such language is Awtuw (Papuan). According to Drossard, "marking for person may only occur with the help of free pronouns. This means that as a rule the author has to gloss the verb forms non-sententially", as in (1):

(1) ka -w -allow -te  
   NG -NF -speak -FUT  
   'can't/won't speak'

In order to get a "sentential" reading a free personal pronoun must be added:

(2) wan ka -w -allow -te  
   Isg NG -NF -speak -FUT  
   'I can't/won't speak'

Boas (1911) shows similar facts for Haida (Na-Dene). Drossard classifies languages of this type as having *non-sentential polysynthesis.* In languages with *sentential polysynthesis,* "an often cited characteristic becomes evident: one word represents one
sentence” (p. 252). Thus, these linguistic types can be distinguished by whether or not there is obligatory person-marking on the verb form.²

Based on evidence from Nivkh (a language isolate of East Asia), Mattissen (2003) re-focuses the discussion of polysynthesis to a more general aspect of morphosyntactic organization that she terms “dependent-head synthesis”, which crucially applies not only to verbs, but to nouns as well. Mattissen defines dependent-head synthesis as

the principle operative for the encoding of possessors, attributes, objects and complement clauses in Nivkh. Dependents (in the sense of Nichols 1986) are generally marked on their heads, i.e. dependents and heads form complex synthetic units. This is true both for objects and complements, marked on their verbs, and for attributes, possessors, and complements, marked on their head nouns. (p. 33)³

The discussion in this chapter will focus on the verbal domain. Although Mattissen raises the intriguing possibility that Comanche has similar properties to Nivkh with respect to complex nouns due to the compounding of “non-root bound morphemes”, “Charney (1993) does not give a description of noun complex formation”, although certain examples in Charney’s grammar “suggest that a noun forms a word unit with its modifiers” (Mattissen 2003: 264). The comparative aspects of the syntax of DPs in Uto-

² Drossard further sub-categorizes languages with sentential polysynthesis on the basis of whether or not these languages have “incorporation” or not, and whether those that have incorporation have “direct” incorporation or incorporation only after “derivation”. The first distinction will be discussed below in my discussion of “pronominal argument” languages, whereas the latter distinction is not made in the Pervasive Syntax Perspective adopted here.

³ Like O’odham, Nivkh does not mark its verbs for subject agreement.
Aztecan, and its relevance to Mattissen's typology of dependent-head synthesis, will be left for future research.

In this chapter we will focus on the technical, syntactic definition of the term *polysynthesis* as used by Mark Baker. This usage correlates to *polypersonalism* as used by Mattissen (2003), and to *sentential polysynthesis* as used by Drossard (1997). The essential idea from this perspective is that a polysynthetic language requires, obligatorily, morphological marking for both the subject and the object on the main verb of a sentence. We ignore here the complex verbs with long strings of adverbs in the "non-sentential" polysynthetic languages (cf. Fortescue 1994, Drossard 1997, Mattissen 2003).

In Baker (1996), the "Polysynthesis Parameter" was defined as one essential feature that he terms the "Morphological Visibility Condition", defined as follows:

(3) Morphological Visibility Condition (MVC) (Baker 1996: 486)

A phrase X is visible for θ-role assignment from a head Y only if it is coindexed with a morpheme in the word containing Y via:

(i) an agreement relationship; or

(ii) a movement relationship.

In the words of MacSwan (1998), "the attraction of the Polysynthesis Parameter is that it appears to account for a number of tightly-associated linguistic characteristics in a single operation. Baker (1996) attributes to this parameter the following properties" (p. 102):
Major characteristics of polysynthetic languages (Baker 1996: 498-9)

- Syntactic noun incorporation (NI)
- Object agreement is obligatory
- Free pro-drop
- Free word order
- No NP reflexive
- No true quantifiers
- Obligatory wh-movement
- N agrees with R argument
- No true determiners
- N agrees with possessor
- Restricted morphocausative
- NI or agreement in PP
- CP arguments only if nominal
- No infinitives

Using evidence from Southeast Puebla Nahuatl, MacSwan (1998) criticizes Baker’s macroparametric approach because of its failure to correctly correlate the macroparameter of polysynthesis with this cluster of properties. MacSwan shows that these correlations fail from both directions—in some cases there are polysynthetic languages that do not have the properties Baker associates with polysynthesis, and in other cases there are non-polysynthetic languages that do.

Just focusing on the evidence from Nahuatl NPs, MacSwan shows that two of the implicational relationships posited by Baker, (4)d and (4)f, do not hold up. In Baker (1996)’s theory, NPs must be adjuncts, and this theoretical point is supposed to explain both their optionality and their free word order.

However, MacSwan claims that NPs are not adjuncts in SE Puebla Nahuatl. One argument in favor of this position is that word order is relatively fixed in Nahuatl, although the basic word order may vary from dialect to dialect. For instance, VSO is the
order most often seen in Classical Nahuatl, and is what has been reported as the most frequent order for the dialects spoken in Huasteca (Brockway and Brockway 1979) and the Malinche Volcano area (Hill and Hill 1986). On the other hand, SVO order has been attributed to the dialects of Tetelcingo (Tuggy 1979), North Puebla (Brockway 1979), Michoacán (Sischo 1979), as well as Southeast Puebla (MacSwan 1998). Although the word order can vary for pragmatic reasons, Nahuatl does not seem to be as flexible in its word-order as other, true “free word order” or “scrambling” languages like Tohono O’odham, Japanese, or Persian. MacSwan also points out that there are plenty of free word order languages that are not polysynthetic, thus the relationship between (4)a-b and (4)d has not been demonstrated.

Second, (Southeast Puebla) Nahuatl also has “true quantifiers” and “true determiners”. According to MacSwan, the NPs modified by such elements must appear in argument positions, “since nonreferential quantified NPs cannot occur in adjunct or dislocated positions” (p. 107) (cf. Rizzi 1986 and Cinque 1990). MacSwan shows that at least one true quantifier, kada ‘each’, has been borrowed into Nahuatl from Spanish cada ‘each’, and this marker exhibits the weak crossover effect expected of a true quantifier but which is lacking in Mohawk (Baker 1996: 57-8).

4 Steele (1976) reports that Classical Nahuatl was in active transition to V-initial order, since it had variable word order. See her paper for a statistical breakdown of word-orders collected from various classical sources.

5 In Baker (1996)’s cross-linguistic consideration of Nahuatl with respect to (4), he found that Classical Nahuatl had positive values for these properties that MacSwan does not find in Southeast Puebla Nahuatl, leading to the possibility that there is diachronic or dialectal variation. However, this fact, in addition to the considerations taken into below, still constitute reasons for casting doubt on the implicational nature of Baker’s macroparametric approach.

6 MacSwan also notes that Classical Nahuatl had a quantifier izquintin ‘each’, “an element which triggers plural morphology on the head noun and the verb”, but “nothing is known about the behavior of Classical izquintin” with regard to weak crossover (p. 109).
MacSwan summarizes his position on the issue of the status of macroparameters within linguistic theory by stating the following:

The conclusions reached here indicate that the basic clause structure of Nahuatl, and perhaps that of other polysynthetic languages, is not radically different from English, Spanish or other well-studied languages. For this reason, while it may be useful for some purposes to refer to polysynthetic languages informally as those languages which are characterized by rich agreement morphology on the verb (sometimes called head-marking languages) plus productive NI, it does not seem likely that polysynthetic languages exist as a formal typological class, contrary to Baker’s proposals. (pp. 112-3)

I agree with MacSwan’s conclusion, and I apply the term “polysynthetic” to languages that exhibit obligatory subject and object agreement plus syntactic NI, i.e. those languages which manifest (4)a and (4)b, as well as an additional stipulation that there be subject agreement on the verb. MacSwan’s view of the structure of Nahuatl as being “not much different” from English is entirely consistent with the account of NI and object polysynthesis developed in Chapter 8, where it is argued that object agreement is the result of NI, and overt object NPs (or DPs) are hyponomous arguments spelling out the lower copy of the moved element. This contrasts with Baker’s theory (and that of Jelinek 1984) wherein full DPs must be adjuncts.

From the perspective I have adopted here, “polysynthetic” can involve either a pronominal element or an incorporated nominal root. This differentiates polysynthesis in Baker’s sense from Jelinek (2001)’s notion of pronominal argument languages, which
require a pronominal element (cf. discussion of Tohono O’odham in 7.4.4, 8.5, and below). True pronominal argument languages lack noun incorporation (Jelinek 2001), thus Nahuatl is a polysynthetic language but not a pronominal argument one. I take this complementary distribution to be a result of an independent parameter setting that requires a pronominal acting as a classifier to be incorporated, rather than a lexical root, in the pronominal argument languages. That is, pronominal argument languages constitute a further level of grammaticalization of noun incorporation, developed from polysynthetic languages.

Note that this aspect of pronominal argument languages is relevant only to object agreement, since subject agreement need not be obligatory on the verb. Tohono O’odham, for instance, requires subject agreement in the second position auxiliary complex, but not on the verb itself (see 7.4.4). Following the terminology of Baker (2001: 148-9), I will refer to the use of agreement morphemes for subjects and objects as subject polysynthesis and object polysynthesis, respectively. If a language employs both, we can regard it as polysynthetic, or having “full-blown polysynthesis”. My adopting Baker’s later terminology should not be taken to be an endorsement of all (or any) of the implicational relationships presented by Baker (1996), given in (4). In the rest of this chapter I will show that many of the properties attributed to Nahuatl’s polysynthetic status can be attributed to Uto-Aztecan more generally. First, however, we consider the evidence for separating subject and object polysynthesis, in order to illustrate how their relative order can be used as evidence for the order of their diachronic fusion, in a process typical of grammaticalization.
9.3. Subject and object polysynthesis in cross-linguistic perspective

The prototypical polysynthetic language has pronominal marking for both subjects and
direct objects. In some cases, e.g. Nahuatl, these bound morphemes are clearly
historically derived from free pronouns. Other languages may have more fusion of the
agreement elements. For example, in the Iroquoian languages subject and object
agreement markers appear as portmanteau morphemes. According to Mithun (1991),
on these elements can be traced back to the Proto-Iroquoian pronominals. Mithun
states that “the Proto-Iroquoian pronominal system already distinguished three persons.
Cognates of first, second, and third person markers can be traced throughout the family,
although different gender distinctions have developed in the different branches” (p. 88).

For example, Seneca has 60 pronominal prefixes, each of which indicate “person,
gender, number, and case of the core participants in an event or state” (Chafe 1996: 561).
Chafe’s reconstructed forms for first and second person passive forms (i.e. the prefixes
that indicate first and second person patients without agents) are given in (5) (gender is
only relevant to third person forms):

<table>
<thead>
<tr>
<th>Patients</th>
<th>1st sg</th>
<th>1st dual</th>
<th>1st pl</th>
<th>2nd sg</th>
<th>2nd dual</th>
<th>2nd plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>(no agent)</td>
<td>*wak-</td>
<td>*yokni-</td>
<td>*yokwa-</td>
<td>*sa-</td>
<td>*sni-</td>
<td>*swa-</td>
</tr>
</tbody>
</table>

Although in the following discussion I follow the Uto-Aztecanist tradition of separating subject and object
markers in Nahuatl, Svenonius (1993) discusses a variety of Nahuat (without the /l/ phoneme) which has
developed a more portmanteau-based paradigmatic system. See his paper for a Lexicalist analysis of this
variety, spoken in San Miguel Tenextatiloyan, near Puebla.
Following Mithun (1991), I take it to be the case that the order of the agreement morphemes on the verb reflects the order in which these elements were grammaticalized, in addition to reflecting the synchronic syntax of the language in question. This position presupposes that languages can develop a stable system with either subject or object agreement, having one in the absence of the other. This is the case.

Baker (2001) presents Chichewa (Bantu) as a language that only has obligatory subject agreement. In Chichewa, subject NPs are totally optional. Object NPs can be deleted if there is object agreement (i.e. incorporation). However, if an object NP does appear without this agreement it has to appear immediately after the verb. With object agreement, the word order of both the subject and object NPs is completely flexible. As Baker puts it, “Chichewa thus combines Mohawk-like and English-like properties. It seems to have a Mohawk-like mode with object agreement and a second mode (without object agreement) in which it treats its direct object essentially as English does” (2001: 146). Mchombo (1998: 500-1)) suggests that this may be a property of Bantu languages more generally. Other languages like this are attested, including Slave (Northern Athabaskan), which has an SOV order when object agreement is not present.

According to Baker,

It is important to realize that not every compound language one can readily imagine constitutes a possible human language. Chichewa and Slave are alike in that they both treat their subjects in the Mohawk way and their objects in the English/Japanese way. One can just as well imagine the opposite blend. This would be a language in which the verb always agreed with the object, so that the
object could be omitted or put on either edge of the sentence. But the verb would not have to agree with the subject, which would have to be present and in its usual position before the main verb or tense auxiliary. (2001: 147)

Baker calls such a language “Reverse Chichewa”, which he concedes “might also allow subject agreement as an option, and when it was used the subject noun phrase could be left out and word order would be freer.” However, “no such language is known to exist” (p. 147).

To the contrary, however, as we saw in 7.4.4., such a language indeed does exist, Tohono O’odham, with one wrinkle. In O’odham, subject agreement is also obligatory, although it does not appear on the verb itself. In fact, subject agreement is cliticized to the AUX element which appears in second position.\(^8\) According to Steele (1977)’s account, what was once a subject pronominal has fused with the tense/aspect AUX element in exactly the position that Baker suggests that subjects should appear, “before the main verb or tense auxiliary”. In addition, other Uto-Aztecan languages have optional object polysynthesis without subject agreement, e.g. Hopi (with obligatory

---

\(^8\) The closely related Tepiman language Pima Bajo may be closer to a “Reverse Chichewa” than O’odham, since it does not have obligatory subject marking in an auxiliary complex. However, it is not clear from Estrada’s (1996) grammatical sketch whether or not object marking is in fact obligatory. If so, this marking is null for third person in both the singular and the plural; however, all verbs with first and second person objects given by Estrada (1996) and Estrada and Steele (1999) are marked for the object. In addition, however, Estrada and Steele give examples of intransitive verbs marked for subjects, as in (i):

(i) \(\text{aan oob nokiaq \text{in-\text{ild}}} \)  
1s Pima speak-fut 1s-want \(\text{Is Pima speak-fut Is-want} \)  
'I want to speak Pima'  

Nivkh (aka Gilyak), a language isolate of Russia, may be another example of a “Reverse Chichewa”. According to Mattissen (2003), “one undergoer is obligatorily marked on the verb, either via dependent-head synthesis of the undergoer noun and the verb or via undergoer prefixation” (p. 203). Subjects (“actors”) are only marked under certain conditions.
object agreement on what K. Hill 2003 calls “pronominal verbs”—see 7.4.4) and Yaqui (7.4.5).

Baker raises the possibility that the polysynthesis parameter should be subdivided into two more limited parameters:

(6) **The Subject Polysynthesis Parameter** (Baker 2001: 148)
    The subject of a verb must be expressed in that verb (e.g. Mohawk, Chichewa, Slave).

(7) **The Object Polysynthesis Parameter** (Baker 2001: 149)
    The object of a verb must be expressed inside that verb (e.g. Mohawk).

Baker rejects this possibility on the grounds that no language is bound by (7) only, since Mohawk- (and Nahuatl)-type (i.e. polysynthetic) languages have (6) as well. However, if we accept that O’odham is a language with (7) applied without (6), as it appears that we must, we have reason to suppose that this proposal is in fact necessary to differentiate Chichewa/Slave-style languages from O’odham-style languages. Mohawk- and Nahuatl-type languages have both, presumably having developed from adding (7) before (6).

The crucial missing piece in a diachronic account that would explain the variable grammaticalization of subject and object agreement morphology would be the observation that different language families could “morphologize” these agreement patterns in different temporal orders—subject first, then object, vs. object first, then subject. In fact, we do see differences along these lines.9 Nahuatl patterns with Mohawk and most polysynthetic languages in having its subject pronominals precede object pronouns; see 9.4 below. In Athabaskan and Cupeño (Takic), however, the order of

---

9 This approach assumes that the order of morphological elements reflects order of morphologization (grammaticalization) rather than the original word order of free pronouns in the proto-system; see Steele (1977), Mithun (1993), and much work in grammaticalization theory.
pronominal marking is the reverse. We will discuss Athabaskan and Cupeno
pronominals in sections 9.3.1 and 9.3.2, respectively.

9.3.1. Pronominal order in Athabaskan

The Athabaskan languages display templatic morphology, wherein specific slots of the
"verb-sentence" are filled by different functional and lexical material. This template
consists of a verb stem and classifier, in addition to a variety of other functional elements
and markers for subjects and objects. Rice (2000) shows that uniformly across
Athabaskan the order of these elements involves subjects being closer to the verb root
than objects, thus yielding templates with O-S-V ordering. Third person subject markers
never appear in the same position as the first and second person subjects, however. Third
person subject agreement typically appears immediately after the object marker, whereas
"first/second person forms are at the right edge of the functional complex" (Rice 2000:
180), with several slots intervening between these two positions. Hoijer (1971)’s proposal
for the Pan-Athabaskan verb template is given in (8):

adverbial(s) + iterative + pluralizing + object + deictic subject + adverbial +
mode/tense/aspect + subject pronoun + classifier + stem

According to Rice, “first/second person subject inflection represents agreement,
including features of person, number, and gender, while third person subject inflection
represents number and gender, but not person” (p. 181). Rice distinguishes between
these two inflectional positions by referring to the first as “Agreement” and the second
“Number”.
Although Rice regards these elements as “agreement” rather than as “pronominal arguments”, the point at issue here is the relative order of the grammaticalization, or morphological fusion, of these elements into the verb template. According to Mithun (1991), it is typical for pronominal paradigms to become bound at different rates, with first and second person typically being bound before third person in most languages, if third person pronominals become bound at all. Similarly, number is not always distinguished at the same time that pronominal elements become bound, and, in Athabaskan different languages have grammaticalized third person number in different ways (Mithun 1991: 90-2).

Presumably, the different positions of these elements in the Athabaskan verb-complex reflect the diachronic order of their morphological fusion. By this I mean that the chronological order of the fixing of these pronominal elements as morphemes obligatorily bound to the verb is reflected in their synchronic order. However, although we can see this from a diachronic perspective, there still must be some synchronic account of the placement of these morphemes, since each generation of the learners of a language must create a new grammar from their surrounding linguistic input. The order of pronominal elements in Athabaskan pose an interesting puzzle for synchronic syntactic approaches to word-formation, such as Baker (1985)’s Mirror Principle, since the subject morphemes appear to be closer to the verb root than the object morphemes.

A secondary issue of interest here is the status of the object marker in the Athabaskan languages. Accepting Rice’s proposal that subject inflection is agreement, as I assume is the case for Nahuatl and other languages with subject-marking on the verb (because of
the inability of agents to incorporate by head movement), what is the status of the object agreement? Jelinek (1989) shows that noun incorporation with open-class nominal roots is a characteristic of Athabaskan generally, but it does not occur in Southern Athabaskan (Apachean), which includes Navajo. These languages are pronominal argument languages, where the incorporated direct object must be a pronominal element. Although the pursuit of this point goes beyond the scope of our discussion here, I assume that an O'odham-like account (see 8.5) could be applied to the Apachean languages, whereas the other Athabaskan languages might have standard noun incorporation. The major theoretical question, though, is the appropriate way to account for the order of the object vis-à-vis the subjects—see Rice (2000) and Hale (2003) for possible accounts. Since Cupeño has similar ordering of pronominal elements, we will focus our attention on an analysis of that language instead.

9.3.2. Pronominal order in Cupeño

Recall from 7.4.2 that the relative order of pronominal elements in Cupeño is similar to what we see in Athabaskan. Verbs in the past tense obligatorily take subject prefixes; it is the use of these prefixes that mark the past tense. The subject prefix paradigm is given in (9):

(9) **Cupeño Subject pronominal affixes** (J. Hill 2003b: 210)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ne-</td>
<td>chem-</td>
</tr>
<tr>
<td>2</td>
<td>'e-</td>
<td>'em-</td>
</tr>
<tr>
<td>3</td>
<td>pe-</td>
<td>pem-</td>
</tr>
</tbody>
</table>

The use of the third person plural prefix for past tense is shown in (10)a, contrasted with
(10)b, which has a future tense verb form and therefore does not have a prefix:\(^\text{10}\)

\[(10)\text{a. } \text{tuku= 'ep pem-chix} \quad \text{(J. Hill 2003b: 210 [6a])}
\]
\[
\text{yesterday=R} \quad \text{3.PL-die.PL}
\]

‘Yesterday they died’

\[
\text{b. tukumdy m=el=pe chix} \quad \text{(J. Hill 2003b: 210 [6b])}
\]
\[
\text{tomorrow} \quad \text{PL=3.PL.ABS=IRR} \quad \text{die.PL}
\]

‘Tomorrow they will die’

Object prefixes are optional for verbs in \textit{any} tense. Hill analyzes these forms as clitics; see discussion in 7.4.2. The object clitic paradigm is given in (11).

\[(11) \text{Cupeno Object pronominal clitics (J. Hill 2003b: 210)}
\]

\[
\begin{array}{c|c}
\text{SG} & \text{PL} \\
\hline
1 & ni- & chimi- \\
2 & 'i- & 'imi- \\
3 & pi- & mi- \\
\end{array}
\]

The use of the object clitic with a past tense form is given in (12); here we see that the object clitic appears before the subject prefix:

\[(12) \text{tukup= 'ep 'i-ché'-max} \quad \text{(J. Hill 2003b: 211 [7])}
\]
\[
\text{yesterday-R} \quad \text{2.SG.OBJ-1.PL-give}
\]

‘yesterday we gave it to you’ (2 43 480)

Under J. Hill (2003b)’s analysis of these constructions, the subject agreement morphology is inserted after the verb head-raises to AgrS, via T, Asp, and \(v\) (Voice). Hill analyzes the object clitic as originating in the “Focus complex” (Kiss 1995), and then

\(^{10}\) Subject pronouns are prefixed to the verb root only in the zero class. There are also two large classes, the \(\text{-in}\) and \(\text{-yax}\) classes, wherein the subject pronoun appears between the verb root and these suffixal thematic elements:

\[(i) \text{yit-ne-n} \quad \text{raise-1.SG-IN} \quad \text{(Barragan 2003: 143 [5], [6])}
\[
(ii) \text{hét-pe-yax} \quad \text{crouch-3.SG-YAX}
\]

‘I raised’ ‘He crouched’

See Barragan (2003) for a DM analysis of these constructions, where he proposes that the thematic affixes are located in \(v\), and the shifting of the subject is claimed to be the result of Local Dislocation at PF.
cliticizing to the verb complex, a "peculiarity" that is "specific to Cupéño" within Uto-Aztecan (p. 224). This is unlike Nahuatl, which presumably has its direct object position base-generated within the VP itself, and which has the expected order of subject preceding the object. It is possible that a clitic analysis might also be available for the objects in Athabaskan as well, although the pursuit of this idea goes beyond the scope of our discussion here.

What is important for our purposes is that subject agreement has grammaticalized in Cupéño, a language that requires this agreement in the past tense. In Baker (2001)'s terms, Cupéño has thus developed the "subject polysynthesis" parameter for verbs in the past tense.

Object prefixes are possible in this language, although since they are not required no "object polysynthesis" parameter has been set. Crucially, though, this variation shows that there are two independent parameters for agreement marking on a verb. Presumably, a future generation of Cupéño speakers could reanalyze the object clitics as obligatory, and if this were to occur then the addition of the object polysynthesis parameter to the subject polysynthesis parameter would yield something like Baker’s Polysynthesis Parameter. However, Cupéño does not have syntactic noun incorporation, thus making it somewhat different from the typical polysynthetic languages that we see elsewhere.

Having justified the distinction between subject and object polysynthesis, we now turn to a brief discussion of the structure of Nahuatl (9.4) before establishing the Uto-Aztecan bases for Nahuatl’s polysynthetic characteristics (9.5).
9.4 Structure of Nahuatl

Most of the varieties of Nahuatl, including Classical Aztec, are unambiguously polysynthetic in Baker (1996, 2001)’s syntactic sense: verbs in this language must be inflected for both subject and object agreement, which in this language take the form of prefixes. The relevant prefixes for Classical Nahuatl are given in (13) and (14), respectively:

(13) **Classical Nahuatl Subject Prefixes** (Newman 1967: 193)

<table>
<thead>
<tr>
<th></th>
<th>Sg.</th>
<th>Pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person:</td>
<td>ni-</td>
<td>ti-</td>
</tr>
<tr>
<td>2nd person:</td>
<td>ti-</td>
<td>am-</td>
</tr>
<tr>
<td>3rd person:</td>
<td>ø-</td>
<td>ø-</td>
</tr>
</tbody>
</table>

(14) **Classical Nahuatl Object Prefixes** (Newman 1967: 193)

<table>
<thead>
<tr>
<th></th>
<th>Sg.</th>
<th>Pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person:</td>
<td>nech-</td>
<td>tech-</td>
</tr>
<tr>
<td>2nd person:</td>
<td>mitz-</td>
<td>amech-</td>
</tr>
<tr>
<td>3rd person:</td>
<td>k(i)-</td>
<td>k(i)-</td>
</tr>
</tbody>
</table>

The reduced forms of the subject pronominals are obviously related to the independent pronouns:

(15) **Classical Aztec Independent Pronouns** (Steele 1977: 571 [15])

<table>
<thead>
<tr>
<th></th>
<th>Sg.</th>
<th>Pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person:</td>
<td>ne?wātl</td>
<td>te?wān(tin)</td>
</tr>
<tr>
<td>2nd person:</td>
<td>te?wātl</td>
<td>ame?wān(tin)</td>
</tr>
<tr>
<td>3rd person:</td>
<td>ye?wātl</td>
<td>ye?wān(tin)</td>
</tr>
</tbody>
</table>

In the long history of contact with Spanish some varieties of Nahuatl are apparently losing some of Classical Nahuatl’s polysynthetic characteristics and taking on a more

---

11 All Nahuatl orthography in this section, which has been collected from various sources from different periods, has been standardized with respect to vowels: a macron indicates a long vowel. In this use I follow the practice of Andrews (2003).
analytic nature—see Hill and Hill (1986) and Farfán (2004) for further discussion of these varieties.\footnote{It is interesting to note that loss of polysynthesis is not an inevitable result of language contact. Trask (1998) discusses the relatively low level of grammatical influence of Indo-European languages on Basque, which has retained its ergative syntax and polypersonal verb agreement even after two thousand years of contact with Indo-European languages.}

In addition to marking a verb with one of the direct object prefixes in (14), object-marking can also occur via the incorporation of a nominal root which is ostensibly the direct object of the verb. Merlan (1976) provides a useful discussion of the pragmatics behind such incorporation in Huauhtla Nahuatl (HN). In general, noun incorporation is used to maintain topicality, i.e. definiteness in discourse. As Merlan explains, “In contrast with complete pronominalization in HN, the nature of incorporation is to permit preservation of the lexical properties of NP adjuncts previously introduced, and thus definiteness of discourse reference is maintained” (p. 185). In Merlan’s terminology this applies only to “contextual” or “discourse-determined” NI; there are also cases of what Merlan refers to as “lexical” NI, which is the obligatory use of NI in certain constructions. In the terms used in this dissertation, this “lexical” NI is simply the use of N-V compounding to generate idiomatic meanings, as in the contrast in (16), where the literal interpretation occurs with the non-incorporating structure (16)a, and the idiomatic reading is generated with NI (16)b:

\begin{itemize}
  \item (16)a. \textit{\textipa{\(\varnothing\)-ne\(\text{\`e}\)-maka-\(\text{\`o}\)-k}} \textit{pa\`ahi} \textit{(Merlan 1976: 184)}
  \begin{tabular}{ll}
    & 3SG-1SG-GIVE-PAST-SG medicine  \\
  'He gave me medicine'  \\
  \end{tabular}
  \\
  \item (16)b. \textit{\textipa{\(\varnothing\)-ne\(\text{\`e}\)-pa?-maka-\(\text{\`o}\)-k}} \textit{(Merlan 1976: 184 [4])}
  \begin{tabular}{ll}
    & 3SG-1SG-medicine-GIVE-PAST-SG  \\
  'He doctored me'  \\
  \textit{(lit. 'He medicine-gave me')}  \\
  \end{tabular}
\end{itemize}
In addition to the incorporation of direct objects, unaccusative subjects and instruments may also incorporate in Nahuatl:

(17) *tlá-ä-weci-0-Ø*
    INDEF-water-fall-PRES-SG
    ‘It is raining’

(18) *Yaʔ 0-ki-kocillo-tete ḟki  pansi*
    he  3.SG.S-3.SG.OBJ-knife-cut  bread
    ‘He cut the bread with a knife’

Tuggy (1986) gives a Cognitive Grammar analysis of incorporation, and also gives examples of the incorporation of direct objects, unaccusative subjects, and instruments in Orizaba Nahuatl (ON) and Tetelcingo Nahuatl (TN).

(19) ON
    *kahven-i*
    coffee-drink
    ‘drink coffee’

(20) TN
    *tōnal-kisa*
    sun-emerge
    ‘the sun comes out’

(21) TN
    *lāpis-k’ilowa*
    pencil-write
    ‘write with a pencil’

Tuggy also discusses “active zone” incorporation, which refers to a body part which is “directly involved in the action of the verb” (p.459). These can lead to different construals, or in the terminology employed in Chapter 3, different idiomatic readings. Some involve instrumental or patient readings, e.g. (22) and (23), respectively:
Some, however, are less predictable. For example, with the incorporation of tōč ‘rabbit’ into the verb root mōlla ‘hurl, shoot’, the meaning is not to ‘hurl rabbits’ or ‘shoot rabbits’, but to ‘hunt rabbits’ (Tuggy 1986: 460).

See Chapter 8 for a formal syntactic treatment of various constructions like these in Uto-Aztecan. What is under our focus here is the incorporation of the direct object; instrumental and “subject” incorporation are adverbial in nature and are brought about through Merge (i.e. compounding). Under the analysis presented in Chapter 8, direct object incorporates, including pronominals, are derived from underlying object position (sister to V) through head movement (i.e. incorporation in the sense of Baker 1988). This even applies to polysynthetic verbs with full DP hyponymous arguments, which, following the principle of Late Insertion, are licensed in the “trace” position (lower copy) left over after this head-movement. An example of a hyponymous argument in Huasteca Nahuatl is given in (24):

(24) ki-piya miak tomi (Beller & Beller 1979: 216)
it-have much money
‘He has a lot of money’

In sentences like this, the direct object pronominal acts as a sort of classifier for the external DP. Some languages, such as O’odham and Navajo, only allow for pronominals
in this incorporated object position, indicating ‘it’, ‘this, ‘thing’, and the like. I take this to be a further grammaticalization based on the “object polysynthesis” type.

The use of incorporated nominal roots and object pronominals in such “classificatory noun incorporation” constructions is found throughout the Uto-Aztecan language family, as we saw in Chapter 7. We now turn to a discussion of this and other aspects of polysynthesis in Nahuatl to show the Uto-Aztecan origins of these characteristics.

9.5. Nahuatl in Uto-Aztecan context

The two defining characteristics of a polysynthetic language, according to the definition of Baker (1996), are subject and object agreement on the verb (“polypersonalism”), and productive syntactic noun incorporation. Each of these aspects have been discussed at length with respect to their productivity in Uto-Aztecan. In sections 2.2 and Chapters 6-8, respectively, it was argued that remnants of each of these can be traced back to Proto-Uto-Aztecan. In 9.5.1 and 9.5.2 we will briefly review the evidence for pronominal clitics/affixes and syntactic noun incorporation across Uto-Aztecan.

9.5.1. Polypersonalism

As discussed in 2.2.3, most Uto-Aztecan languages have subject and object clitics of some kind. Some languages have second position clitics, some have pre-verbal clitics or affixes, and some have both. Langacker (1977a) reconstructs subject and object clitics for Proto-Uto-Aztecan. He states that “it seems fairly clear that subject clitics, but not
subject verb prefixes, can be reconstructed for P-UA" (p. 126). He offers the following as speculative reconstructions, "offered mainly as a basis for research and discussion":

(25) **PUA Subject Clitics** (Langacker 1977a: 126)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1P</td>
<td>*(=ni)</td>
<td>*(=ta)</td>
</tr>
<tr>
<td>2P</td>
<td>*(=i)</td>
<td>*(=i-mi)</td>
</tr>
<tr>
<td>3P</td>
<td>*(=pi)</td>
<td>*(=pi-mi)</td>
</tr>
</tbody>
</table>

(26) **PUA Independent Definite Pronouns** (Langacker 1977a: 124)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1P</td>
<td>*(i-)ni</td>
<td>*(i-)ta(-mi)</td>
</tr>
<tr>
<td>2P</td>
<td>i(-mi)</td>
<td>*'i-mi</td>
</tr>
<tr>
<td>3P-H</td>
<td>*pi</td>
<td>*pi-mi</td>
</tr>
<tr>
<td>3P-NH</td>
<td>*a</td>
<td>*a-mi</td>
</tr>
</tbody>
</table>

Following Langacker (1977a), Steele (1978), and much other work, I assume that the pronominal affixes in Nahuatl have developed from these PUA clitics. What is interesting in the context of this discussion is that the cross-linguistic evidence from other Uto-Aztecan languages indicates that this was a gradual process. Having argued that we can differentiate the grammaticalization of obligatory subject agreement and obligatory object agreement, we will consider each of these in turn.

**9.5.1.1. Object pronominals**

I assume that the obligatory object polysynthesis, which we see in Nahuatl, must have developed from *optional* object polysynthesis. This position is supported by the fact that we see other Uto-Aztecan languages with optional pronominal object agreement, in addition to noun incorporation. Optional object clitics in Cupeño were discussed above.
We also see optional pronominal agreement for third person in Yaqui:

(27) \textit{Inepo Hose-ta (aa)-vicha-k} \hfill (Maria Amarillas, p.c.)
\begin{tabular}{l}
1.SG \text{Hose-ACC} \hfill (3.SG.\text{OBJ})-\text{see-PERF} \\
\end{tabular}
'I saw Hose'

(28) \textit{Huan Hose-ta into Maria-ta (am)-vicha-k} \hfill (Maria Amarillas, p.c.)
\begin{tabular}{l}
Huan \text{Hose-ACC} \hfill \text{Maria-ACC} \hfill (3.PL.\text{OBJ})-\text{see-PERF} \\
\end{tabular}
'Huan saw Hose and Maria'

Additionally, Hopi has a small class of verbs, the "pronominal verbs", that require object agreement (K. Hill 2003):

\begin{tabular}{l}
little.boy older.brother 3\text{rd}.SG.-keep.wanting.to.hang.around.with \\
\end{tabular}
'The little boy wants to hang around with his older brother.'

b. \textit{Nu’ nuunukpàntuy amìu-pinya} \\
\begin{tabular}{l}
1sg evil.ones-ACC 3\text{rd}.PL-make.leave \\
\end{tabular}
'I made the bad guys go away.'

I propose that PUA had similar patterns of optional pronominal object marking. Over time, Pre-Nahuatl (PN) grammaticalized obligatory object agreement, for all transitive verbs that did not have overt noun incorporation. That is, at some point PN set the "object polysynthesis parameter", and at this stage PN was like Tohono O'odham, which has obligatory object agreement (7.4.4). At this stage this parameter could be satisfied by either the incorporation of a nominal root or the incorporation of a direct object pronomial. By hypothesis, at this stage subject agreement was not obligatory.
9.5.1.2. Subject pronominals

It appears that PUA had subject clitics that could either be placed in the second position or prefixed to the verb. Several Uto-Aztecan languages retain this feature, including Yaqui, where these clitics are optional and can appear in either position:

(30) *kwarentapeso dydryota-ne ne-koba iani inine* (Steele 1977: 543 [9])
    forty peso daily=CP CP=earn now here
    'Now I make forty pesos a day here'

Some languages only have second position clitics, and some only have pre-verbal clitics (or prefixes). Although Steele (1978) argues that pre-verbal clitics are secondary developments from the second position clitics, it is simpler to suggest that some Uto-Aztecan languages have simply “run with” one or the other; see discussion in 2.2.3.

I suggest that it was after FN already had obligatory object polysynthesis that these pre-verbal subject clitics were reanalyzed as obligatory subject agreement, and it was then that they “fused” to (i.e. became morphologically a part of) the verbal complex. It was with this reanalysis that PN then had obligatory subject agreement, and thus had set the “subject polysynthesis parameter.”

9.5.2. Syntactic noun incorporation

The second crucial element to (4) is syntactic noun incorporation. Noun incorporation in Uto-Aztecan has been discussed at length in this dissertation; see especially Chapter 7 for the comparative aspects of NI in Uto-Aztecan. I have proposed that Proto-Uto-Aztecan had both syntactic NI and classificatory NI.
The crucial point to be made here is that noun incorporation is independent of the Polysynthesis Parameter, since non-polysynthetic languages also have syntactic NI (e.g. Hopi). The one correlation that I will make here between syntactic NI and object polysynthesis is that it is my hypothesis that object polysynthesis derives from classificatory NI. Further, the Pronominal Argument Parameter, as defined by Jelinek (2001), is a further step in grammaticalization, where NI of nominal roots is no longer possible, but instead transitive verbs require the insertion of classifying elements into the highest position of the NI chain.

This grammaticalization cline is shown below in Table 9.1:

<table>
<thead>
<tr>
<th>Syntactic and Classificatory NI</th>
<th>Optional Object Polysynthesis</th>
<th>Obligatory Object Polysynthesis</th>
<th>Pronominal Objects Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporation of direct object nominals and pronouns; hyponymous objects</td>
<td>Optional pronominal agreement markers</td>
<td>Incorporation is required: either pronominal or nominal</td>
<td>No nominal incorporation; Pronominals (or classifiers) only</td>
</tr>
<tr>
<td>Hopi</td>
<td>Pre-Nahuatl; Yaqui</td>
<td>Nahuatl</td>
<td>Tohono O’odham</td>
</tr>
</tbody>
</table>

*Table 9.1. Grammaticalization trajectory of (object) polysynthesis*

A crucial aspect of this analysis is that subject agreement is completely independent of the object incorporation referred to in Table 9.1. Although Nahuatl has developed obligatory subject agreement, languages like Tohono O’odham (and other Tepiman languages) show us that this is not a necessary development.
9.5.3. Related issues

There are two further aspects of grammar that have been linked to polysynthesis that we can address with respect to the Uto-Aztecan languages. These are discourse reference for incorporated nouns (9.5.3.1) and discontinuous constituents (9.5.3.2).

9.5.3.1. Discourse reference for incorporated nouns

The discourse function of incorporated nouns in Nahuatl has been addressed by Merlan (1976), as discussed in 9.4. In Huautla Nahuatl, incorporated nouns can have definite reference. According to K. Hill (2003), discourse reference is also active for incorporated nominals in denominal verb and noun incorporation constructions in Hopi:

(31) *Nu' pakiw-maqto-ni;
   I fish-go.hunting-FUT;

   noqw itam pu-t enang nōōnōsa-ni.
   so we that-ACC in.addition.to eat(PL)-FUT
   'I’m going fishing, so we can eat it (fish) along with the other food.'

It is unclear at this point what the status of discourse reference is for incorporated nominals across Uto-Aztecan generally. With discourse reference for incorporated nominals in individual SUA and NUA languages, however, there is a possibility that PUA had it as well. However, more thoroughgoing investigation into a wider sample of languages would be required to make this a firm conclusion.

Ariel (2001) stresses that there is a continuum of reference for bound pronominal elements crosslinguistically, with some languages allowing for anaphoric reference for bound pronouns and others not. This observation seems to apply to incorporated nouns.

---

13 K. Hill (2003) states that in Classical Nahuatl "incorporated nouns seem to be only indefinite" (p. 241).
as well, providing another argument that there is an affinity between the two, or, from a negative perspective, at least showing that referentiality cannot be used as an argument for distinguishing between the two. Thus, it would seem highly possible that even closely-related languages within a family could vary on this point, as Classical Nahuatl and Huautla Nahuatl apparently do. We conclude that discourse reference for incorporated nouns cannot be used as a criterion for polysynthesis.

9.5.3.2. Discontinuous constituents

J. Hill (2003a) provides a recent study of extensive discontinuous constituency in Cupeño, a non-polysynthetic language of the Takic sub-branch of Northern Uto-Aztecan. Hill shows that a wide range of discontinuity can occur in nominal constructions in that language, even though nominals are in argument positions. This poses a conundrum for Baker (1996)'s account of discontinuity, which relies on the adjunct status of NPs with discontinuous elements. However, with the exception of numeral modifiers, according to Baker (1996: 141) discontinuous constituents “are either outright impossible or so rare and restricted that they go unobserved in standard grammars” of several languages, including Classical Nahuatl. According to Hill, however, discontinuous constituents have been described for Nahuatl, as well as for Tohono O’odham and Luiseño. Thus, polysynthesis is not a prerequisite for discontinuous constituents either.
9.5.4. Summary

To sum up this chapter so far, I have argued for the following stages in the gradual development of polysynthesis in Nahuatl. First, Pre-Nahuatl developed from a proto-language that shared several properties with Proto-Uto-Aztecan: syntactic and classificatory noun incorporation, as well as subject and object clitics. These have been retained in most Uto-Aztecan languages.

Subsequently, PN developed obligatory object polysynthesis, a requirement that transitive verbs incorporate their complements. This is still the case in modern Nahuatl. By hypothesis, however, at this stage PN did not require subject agreement on the verbal complex. A further development from this stage is the Pronominal (Object) Argument parameter, wherein only pronominals (or other classificatory elements) can be inserted at the highest copy (e.g. Tohono O'odham).

Finally, “full-blown polysynthesis” developed when subject clitics “fused” to the O-V complex, becoming a mandatory requirement on all verbs. Thus, the setting of the Subject Polysynthesis Parameter and Object Polysynthesis Parameter were separate events. While the over-all development of polysynthesis was gradual in nature, the fixing of object incorporation and subject agreement as obligatory could have been “catastrophic”, occurring in a single generation of speakers who reanalyzed previously optional processes as obligatory, thus setting the parameters in question.
9.6. Conclusion

In this chapter I have shown that multiple non-polysynthetic Uto-Aztecan languages display properties usually attributed to polysynthesis, including discourse reference for incorporated nouns, classificatory noun incorporation (i.e. incorporation of a nominal root into a verb stem which then takes an external, hyponomous argument), discontinuous constituents, etc. Each of these indicate that Proto-Nahuatl (PN) emerged from a linguistic system already "susceptible" to the development of polysynthesis, rather than developing this property overnight. That is, languages related to Nahuatl show "precursors" to full-blown polysynthesis (i.e., syntactic and classificatory NI plus subject clitics). The evidence surveyed here is in accord with MacSwan's conclusion that the development of polysynthesis in a given language does not entail Baker (1996)'s implicational relationships (cf. 4).

The proposal made here is consistent with divergent work on grammaticalization and more formal approaches to language change. My approach links a gradual syntactic change, the development of obligatory NI with the fixing of subject clitic-placement, to the abrupt change required by approaches to diachronic syntax focused on child language acquisition (e.g. Lightfoot 1991). If this gradualist approach is correct for the Polysynthesis Parameter, then it could probably also be applied to other proposed "macroparameters" as well. Thus, this line of research suggests the need to search for potential "precursors", in the form of retentions, in other languages and language families with no written record, in order to find (and reconstruct) likely sources for the "grammaticalization" of particular "macroparameters".
REFERENCES


——. In prep. Denominal verbs in Uto-Aztecan. Ms, University of Arizona.


presented at the Utrecht Workshop on Agreement, 2001).


Kennedy, Robert. 2001. The role of markedness constraints in linguistic variation and change. Paper presented at the 6th Southwest Workshop on Optimality Theory (SWOT VI), University of Southern California.


---. 1997b. Cat as a phrasal idiom. Ms, MIT.


---. 1993. Prosodic Morphology I: Constraint interaction and satisfaction. Ms, University of Massachusetts—Amherst and Rutgers University.


Academy of Arts and Sciences: 1-296.


