FEATURES, POSITIONS AND AFFIXES IN AUTONOMOUS MORPHOLOGICAL STRUCTURE

by

ROBERT ROLF NOYER

A.B. in Linguistics with Other Subjects (Greek), Harvard College
(1988)

Submitted to the Department of
Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 1992

© R. Rolf Noyer, 1992. All rights reserved

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

Signature of author ______________________________

Department of Linguistics and Philosophy
August, 1992

Certified by ______________________________

Professor Morris Halle
Thesis Supervisor

Accepted by ______________________________

Professor Wayne O'Neil
Chairman, Departmental Committee

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
OCT 02 1992
LIBRARIES

FEATURES, POSITIONS AND AFFIXES IN AUTONOMOUS MORPHOLOGICAL STRUCTURE

by

ROBERT ROLF NOYER

Submitted to the Department of Linguistics and Philosophy on August 26, 1992 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

This thesis develops a theory of the mapping from the output of syntactic rules to the input to phonology within a component of Morphology. Three types of information are shown to be independently manipulable in the algorithm deriving fully inflected words: morphosyntactic features, phonological material (affixes), and positions-of-exponence.

To capture categorial neutralizations at the morphological level, a filter-based approach to the Impoverishment of morphosyntactic representations is proposed. Filters are argued to be automatically assumed unless positive evidence suppresses them in the learner. Neutralizations are also shown to occur at three other levels: (1) underspecification of realization rules/affixes, (2) disjunction of affixes competing for a position-of-exponence, and (3) transparency of redundant morphosyntactic values. A hierarchy of features is proposed, which, in the unmarked instance, governs the direction of categorial neutralization and the ordering of affixes and morphological rules, where not determined by syntax or by the principle of descending complexity due to Panini.

The mapping to phonological form is proposed to take place in two phases. In the first phase, morphosyntactic well-formedness is established through the Impoverishment of morphosyntactic features and the Linearization, Rebracketing, Merger (Marantz 1988) and Fusion of syntactic heads, giving a level of constituents termed M^0s. In the second phase, the M^0 string is mapped to the well-formed morphological word through rules and affixes. Morphological words are shown to display idiosyncratic non-phonological conditions of well-formedness (obligatory or limited positions-of-exponence) which comprise *autonomous morphological structure*. Particular attention is paid to cases of M^0 splitting, where features from a single morphosyntactic source are realized at multiple positions.

Features in input morphosyntactic representations are argued to be *discharged* through the instantiation of affixes, sometimes at obligatory positions-of-exponence, which are also then discharged. This discharge constitutes spell-out or realization of phonological material. Because both features and positions may be discharged independently, it becomes possible to express disjunctions across position classes, termed here *discontinuous bleeding*. Further, the Rule-Affix Hypothesis states that rules change phonological material and do not discharge morphosyntactic features; affixes add phonological material and do discharge such features.

The Introduction situates the current proposal within prevailing theories of inflection, Lieber (1980, 1992), Anderson (1981, 1992), and Halle (1989a, 1992). Chapter 1 exemplifies the theory through a detailed analysis of the Afroasiatic prefix-conjugation. Chapter 2 gives an explicit theory of person-number features and their interpretation, in particular the inherent number system of Kiowa-Tanoan. Chapter 3 treats complex multiple argument agreement systems with particular attention to the transitive clitic sequences of Nunggubuyu and Ket.

Thesis Supervisor: Dr. Morris Halle
Title: Institute Professor

ACKNOWLEDGMENTS

Striving to earn Morris Halle's respect has been more instructive than any book, any article, any course. All who have worked with him or under his direction know his powerful and profound influence and inspiration: as a third generation of students begin to enter his tutelage, I am honored to have numbered among these. He let no idea, new or old, pass without a challenge; to live is to argue, to argue is to learn. We have done all three.

To work with Ken Hale has always been a great pleasure. His commitment to minority language communities everywhere and to linguistic theory should be a model for all who call themselves linguists. Ken opened up to me the world of Kiowa-Tanoan by sharing with me his work on Jemez; my only regret is that I discovered this world so late in my stay at MIT. As is obvious, this thesis depends a lot on data from little-studied languages. With Ken, I came to appreciate that each of these is indeed an elusive and precious treasure.

Michael Kenstowicz never let me forget that I am (still and always) a phonologist, and what a comfort it was to have him around to keep that alive in me! If Michael had not suggested that I look at Modern Arabic dialects, chapter 1 would never have come about. Thinking about each little problem that Michael drew my attention to ultimately reshaped much of what this thesis contains.

Alec Marantz arrived at MIT in my third year and I knew "things were going to change around here." When Morris and I had reached an impasse, Alec jump-started the morphology engine and away we went. No matter how unintelligible I became, Alec always seemed to figure out my ideas instantly, and would zoom beyond my latest thought to a new one. With Alec I got the sense that there has always got to be a better theory and that I am contributing to it, right now. Why else are we in this business? And how else could one ever finish a thesis?

I also wish to thank my other teachers of linguistics, both old and new: Noam Chomsky, Irene Heim, David Pesetsky, Jim Harris, Ken Wexler, Jim Higginbotham, Richard Larson, François Dell, Wayne O'Neil, Andrea Calabrese, Karin Michelson, Jill Carrier, John Whitman, Susumu Kuno, and Cal Watkins.

Among all present and former students and members of the MIT community, who have in countless ways offered assistance and made life possible, I would like especially to thank Bill Idsardi, Kumiko Murasugi, Doug Jones, Itziar Laka, Sabine Iatridou, Eulàlia Bonet, Michael Hegarty, Hamida Demirdache, Diana Cresti, and in particular Utpal Lahiri, who was always there when I needed him this last year.

I also wish to acknowledge friends outside of the linguistics world, without whom I'd never have gotten through my eight years in Cambridge: Hajime Tokuno, Bill Widnall, Kenro Kusumi, Susan Schwartz, Jay Dickson and especially Michael Majeski, who taught me much about life and happiness.

This thesis could never have been written without the work of many scholars who, in the past 35 years, continued to do primary descriptive and analytical work on little-described languages. To these people, most of whom I have not met, I owe a special debt. It is my hope that what insights this work may provide will enrich their future work as much as theirs has made mine possible:

Ernest T. Abdel-Massih (Tamazight Berber)
Barbara Allen, Don Frantz, Donna Gardiner (Southern Tiwa)
Ben Elson (Sierra Popoluca)
Nora England (Mam)
Eugene Buechel (Sioux)
A.P. Dul'zon and E. A. Krejnovič (Ket)
Ken Hale (Jemez, Hopi, Warlpiri, Winnebago)
Jeff Heath (Nunggubuyu)
D. Hutchisson (Sursurunga)
Eloise Jelinek (Lummi)
Paul Kroskrity (Arizona Tewa)
Gordon McKay (Rembarrnga)
F. Nordbustad (Iraqw)
K.-G. Prasse (Tuareg)
Shlomo Raz (Tigre)
Randall Speirs (Rio Grande Tewa)
Robert Sprott (Jemez)
Emily Stairs and Elena Hollenbach (Huave)
Jorge Suárez (Huave)
Jan Voorhoeve (Bamileke)
Laurel Watkins (Kiowa)
Ursula Wiesemann (ɣɔmálá?)
J.-C. Zeltner & Henri Tourneux (Chadian Arabic)

and many others

Finally, I thank my sisters and most of all, my parents. For a childhood of ideas and books, for the means to make it in this world, for constant encouragement and love, I thank them.

This thesis is dedicated to the memory of my grandfather, Prof. Ralph W. Noyer, a scholar of the Classics, whose spirit of erudition inspired me to the academic life.

Rolf Noyer
August, 1992

Table of Contents

CHAPTER 3: Problems in Multiple Argument Agreement

0. Prefatory Remarks

> Au sein de la morphologie nominale, la flexion hétéroclitique reste un corps étranger. On la décrit, on en signale çà et là les vestiges: on ne l'explique pas.
>
> Benveniste 1935: 4-5

The challenge of any linguistic theory is to separate the contingent, accidental phenomena of language from those principled phenomena which reflect traits intrinsic to the human cognitive apparatus. For this reason, I begin this thesis with Benveniste's astute delimitation of what one can hope to accomplish with a linguistic analysis: in morphology, there will always remain a substantial residue of facts which 'one does not explain,' the *corps étranger* within the heart of any inflectional system. Nevertheless, the goal of linguistic science of whatever persuasion is to continually erode and redefine this residue. As theories of syntax become more *minimalist* in scope (Chomsky 1992), the burden on the morphological component to encompass the accidental, the 'peripheral PF-options' of a grammar, will necessarily increase. Given the immense complexity of inflectional systems, especially those with multiple-argument agreement, one must assume that morphology is more than a residue, more than a lexicon whose contents is a 'prison' which 'contains only the lawless' (DiSciullo & Williams 1987), more than a *corps étranger*.

A morphological theory is, I believe, in large part a means of evaluating a given system of word-formation as more or less deviant from an 'unmarked' state. There are few universal generalizations in inflection, few absolutes. Any model of morphology must therefore be a model of markedness, a more modest theory which predicts the expected rather than the necessary. But such a model may indeed still be an empirically sound one, inasmuch as it defines the expected as that which is automatically expected by children during language acquisition. The present thesis is concerned only with that word-formation which is, broadly construed, inflectional, being realized in the mapping from the output of syntax to the input of phonology. The theory developed here considers three areas of deviation from this hypothetical state of expectations for this mapping.

The first is the construction of inflectional categories. It will be argued that the knowledge of a language's inflectional morphology consists not merely in knowing a list of affixes and minimal principles governing their combination. Rather, I will show that the most explanatory account should include a theory of *filters* which *Impoverish* morphosyntactic representations, preparing them for realization as phonetic sequences. The theory of Impoverishment filters developed here is based in part on Calabrese's (1988,

1992) work on the phonological filters which define the phonological categories, and expands on the original proposal of Impoverishment for Catalan in Bonet (1991). In the present theory, Impoverishment defines the available morphological categories. Filters are argued to be automatically assumed by children unless positive evidence is presented that a given category remains distinct. The arrangement of filters in implicational hierarchies further evaluates the markedness of a given morphosyntactic category.

A simple example of Impoverishment is the following. In French, gender is limited to third person arguments and is available for both singular and plural numbers, hence *il* 'he,' *elle* 'she,' *ils* 'they (m.)' and *elles* 'they (f.).' In German or Russian on the other hand, gender is limited to only the third person singular categories, hence *sie* or они *oní* are unmarked for gender, even though singular nouns invariably have some lexical gender. To express this difference, the filter theory requires that German and Russian contain the filter *[pl gender]. I then argue that the value of the gender feature is deleted at the beginning of a morphological derivation. Gender is deleted because it is lower than number on a universal hierarchy of features, which I argue also orders affixes and morphological rules in the unmarked instance. A purely affix-based theory cannot express such basic generalizations: for such a theory, it is an accident of the lexicon that a given language lacks certain categories, since it is an accident that no affix or stem happens to realize a given logically permissible feature combination.

The second area of deviance is the relation between syntactic atoms, X^0s, and the constituents of morphological analysis, termed here *affixes*. The theory of syntactic affixation, begun with Chomsky (1957) and recently revived in the work of Fabb (1984), Baker (1988), Travis (1984) and Lumsden (1987), has proposed that the constituent structure of complex words is at least in part derived from syntactic movement of X^0s. This has led in turn to the implicit equation of X^0s and affixes, common in much recent work in morphosyntax. In this thesis, while establishing that this equation is not direct, I propose that affixes and X^0s are isomorphic in the *unmarked* instance. Two processes alter strings of X^0s in the mapping to phonological form. These are the Fusion of X^0s into single constituents, called here M^0s, and the possibility of 'splitting' or 'fission' of M^0s into multiple positions. A simple example of fusion is the simultaneous or cumulative realization of Case and Number in the nominal desinences of the more archaic Indo-European languages, as opposed to their separate realization in most Australian and Altaic languages. Splitting is exemplified in chapter 1 in an extensive discussion of the Afro-Asiatic prefix conjugation where agreement features appear in both the prefix and suffix positions-of-exponence.

Splitting of M^0s may be free, in the sense that a given collection of features may be realized as none or many affixes, or splitting may be constrained by language-specific conditions of *autonomous morphological structure*, which define the affixal requirements for a given root to become a well-formed word. These requirements exist separately from requirements of syntax or phonology and motivate a separate level of well-formedness within Morphology. Where splitting is free, there are no obligatory positions-of-exponence required to make a root/stem into a well-formed word, and neither can there be Ø-affixes. Where there are conditions of autonomous morphological structure, required positions, called here Qs, may be realized as garden-variety affixes, contentless (elsewhere) affixes or as Øs in order to fulfill requirements of well-formedness.

The third area of deviance is the relation of position of X^0 to position of the affix(es) realizing the features borne by this X^0. In the unmarked state, the two are identical and each syntactic X^0 corresponds to a phonological string. However, two processes may complicate this direct relationship. The first is Merger as discussed in Marantz (1988), whereby an element may exchange a relation of adjacency with the merger relation to the head of its complement. At the level of syntax, Merger entails head movement in the sense of Travis (1984); at the level of morphology, Merger entails head-affixation. It will be shown that Merger applies crucially in the derivation of Nunggubuyu and Ket verb-words. Second, where there is splitting of M^0s into multiple positions-of-exponence, I argue that affixes align themselves (or, equivalently, the rules which realize the affixes are ordered) according to a hierarchy of features which has a universally least marked character. The hierarchy puts patients above agents, person above number above gender, participants above nonparticipants in the speech act, and first person above second person. Whether this hierarchy is mentally represented as a tree-like graph is unclear but is discussed in the Introduction. Thus, the hierarchy of features governs the direction of Impoverishment and the unmarked ordering of rules/affixes. Morphological rule ordering is proposed to obey the hierarchy of features and the principle of descending complexity, generally attributed to Panini (Kiparsky 1973).

The title of the thesis is intended to emphasize the independence of three types of representations in the mapping from the output of syntax to the input to phonology. These elements are morphosyntactic features, positions-of-exponence within a well-formed word, and affixes, either segmental or autosegmental phonological information. In the traditional structuralist conception of the morpheme, all three elements are bound together: a given morpheme is a relation between sound (an affix), meaning (features), and, implicitly, the position occupied by the sound. Various extensions and alternatives to the classical conception, such as Anderson's Extended Word and Paradigm theory (1981, 1986) and more recently A-Morphous Morphology (1992), the affix-based models of Lieber (1980, 1992) or Williams (1981), and Halle's abstract morphology (1989a, 1990, 1992), have adjusted or modified this basic construct without fully emancipating each of its components. This thesis is devoted to exploring the independence of features, positions and affixes.

The theory I develop permits morphological rules (instantiating affixes) to *discharge* both features from syntax and positions in autonomous morphological structures which define well-formedness at the word level. For example, the Arabic prefix-conjugation will be shown to require a single prefix and a single suffix position of exponence. Morphological rules/affixes discharge the agreement features from syntax in virtue of instantiating affixes at these positions, which positions are in turn discharged as well. Thus, the discharge or 'filling' of a position is independent of the discharge or 'spell-out' of a morphosyntactic feature. When features and positions are invariably associated, as in the classical morpheme model, the independence of features and positions never arises. But where there is M^0 splitting and autonomous conditions of well-formedness, I will show that this independence is crucial, particularly in the expression of affix disjunctions *across* position classes, termed here *discontinuous bleeding.*

To capture the effects of discontinuous bleeding, I take seriously the idea from Carstairs (1987) that each morphosyntactic feature may have a *principal exponent*, essentially its primary inflectional realization. In affix-based theories, the principal exponent of a feature is that affix which percolates the feature; paradigm-based theories, on the other hand, do not generally recognize principal exponence. In the proposed theory, a feature's principal exponent is the affix which discharges the feature, and all other occurrences of that feature in subsequent rules merely condition allomorphy.

This thesis is organized into a long Introduction followed by three additional chapters. The Introduction situates the proposed theory within contemporary conceptions of inflectional morphology, as introduced by Halle (1989a), Lieber (1980) and Anderson (1981). These theories are briefly compared with a fragment of the Latvian nominal declension, which I use to illustrate the *discharge* effect. A comparison of each theory's conception of what a rule is vs. what an affix is leads to an explicit proposal regarding the rule-affix distinction. Essentially, rules are defined here as any morphological process which changes phonological features or morphological features; an affix instantiates phonological information and discharges morphological features and in some cases a position-of-exponence. I then turn to the relation of X^0s to the positions-of-exponence of their features, and then to a theory of Impoverishment.

The first Chapter exemplifies the theory developed in the Introduction through a detailed analysis of the Afroasiatic prefix conjugation. The phenomenon of discontinuous bleeding and the independence of features and positions-of-exponence form the focus of sections 1.0-1.11. In the remainder of the chapter, historical-comparative evidence supports the historical stability of discontinuous bleeding. I also show that certain historical developments can be explained directly as gain or loss of Impoverishment filters, with the inventory of rule/affixes remaining relatively constant. Finally, I exemplify the most radical example of M^0 splitting within Afroasiatic through an analysis of the Berber verbal inflection and small dialect differences within Berber and Tuareg.

Chapter 2 is devoted to developing an explicit theory of person and number features. Such a theory predicts not only what are and are not possible AGR categories in terms of combinations of person-number features, but also how marked a given category will be in terms of the filters which must be suppressed to permit a given category to surface. I argue that the same features figure in phonological interpretation of AGR as inflectional material as figure in the semantic interpretation of AGR categories. In addition, I propose that the 'meaning' of a person-number category has both a truth-conditional and a conversational or functional aspect; only when both are recognized can the iconic relationship of morphosyntactic markedness to morphological markedness be maintained. This is illustrated in particular with an analysis of the functional duals and trials of Nunggubuyu. A variety of theoretical issues are addressed, including the underspecification of person-number features, the so-called 'syou' category of Algonquian, the various means (explicit, functional, inherent, and 'negative') for expression of dual and trial, and the phenomenon of feature switching or α-notation rules in Mam and in the inherent number system of Kiowa-Tanoan.

In Chapter 3 I address problems in very complex multiple-argument agreement systems. I argue that slot- or template-based morphology actually incorporates two separate and distinct elements: the Placing or ordering of affixes and the Licensing of affixes. Recognizing the independence of these two processes allows templates to be understood as more or less marked rule-orderings or affix-hierarchies. I propose that affix order, when not governed directly by syntax, obeys the hierarchy of features in the default case. In addition, Impoverishment is shown to neutralize many distinctions at the featural level before morphosyntactic representations are discharged as affixes. Neutralization and Placing are extensively exemplified in a thorough analysis of the Nunggubuyu transitive clitics. Evidence from the Kiowa-Tanoan languages and Iraqw illustrate the filter-based theory of neutralization. Finally, the process of Morphological Merger is again shown to be at work in placing the INFL constituent within the verb-word in Ket.

0.1 Theoretical Assumptions: Towards a Theory of Morphology after Syntax

The model assumed here differentiates derivational from inflectional morphology in a theory-internal way. Derivational morphology consists of a set of rules and principles operating over a set of roots and derivational affixes to construct the atoms of syntactic structure, denoted here X^0s. The argument structure properties of these units project the d-structure of syntax. I will have little to say about these processes.

Inflectional morphology is a set of rules and principles which supply to X^0s any phonological form whose realization is sensitive to some syntactic process, including agreement (a syntactic indexing relation), concord (a syntactic copying relation), case-marking (morphological realization of various syntactic and thematic relations which arguments bear to predicates), or incorporation (X^0 movement). These processes will be argued to take place after syntax in the mapping to Phonological Form in a component called Morphology. The theory to be defended here is therefore of the 'Weak Lexicalist' variety, by which it is meant that not all word formation occurs prior to syntax.

The derivation of what will be termed here *morphological words* proceeds through several steps. I will assume that every syntactic head (X^0) consists wholly or in part of a structure of morphosyntactic features called a *morphosyntactic representation.* These representations are fully-specified by the end of syntax (Lumsden 1987) but are argued to undergo simplification in the first phase of Morphology. This simplification, motivated by language-specific ceilings on the complexity of inflectional categories (cf. Calabrese 1988), will be called *Impoverishment* (Bonet 1991).

Because linearization is a property of phonological structure and not of syntactic structure, I will assume that X^0s are not linearized in syntax. Linearization of X^0s applies in the first phase of Morphology as well, supplying adjacency relations to all constituents. After linearization, rebracketing under adjacency (Sproat 1985) may apply, creating new constituents within the morphological string. Rebracketing is assumed to be a subtype of the process of morphological Merger (Marantz 1988), which can then create new 'merged' constituents from two syntactic heads by the operation of Fusion. These merged constituents will be called M^0s. This comprises the first phase of Morphology. In the second phase, two demands govern the mapping from strings of M^0s (and their morphosyntactic representations) to phonological strings, which provide the input to Phonology. These demands are (1) well-formedness conditions on morphological words, which I will call *autonomous morphological structure*; and (2) two types of rule: first, rules which change phonological material (readjustments), and second, rules which supply phonological material (affixes) to stems, thereby *discharging* features of the morphosyntactic representation. Each of these processes will be defined and illustrated in the Introduction. The entire process is illustrated below:

(1) The Mapping from Syntax to Phonology

Syntax

↓

Phase I: Morphosyntactic well-formedness:
Linearization, Rebracketing/Merger/Fusion, Impoverishment

↓

M^0 String

Phase II: ↓

Well-formedness conditions on
morphological words --------> morphological word <----- rules / affixes

↓

Phonological Form

We will consider each of these in turn, beginning with a discussion of Phase II.

0.1.1 Well-formedness Conditions on Words

Morphology must in all cases derive well-formed morphological words of a language. For example, many languages require that roots be augmented by 'thematic' stem formatives and supplied further with inflectional affixes. In (2) and (3) two such examples are shown from Latvian (Halle 1992) and Huave (Matthews 1972a, Stairs & Hollenbach 1981):

(2) gulb-i-m 'swan (dat.sg.)'
swan-TH -dat

(3) s - a -niing-an 'our (excl.) house' [1]
1Px-TH-house-augPx

[1]Px denotes 'Possessor Affix', and 'aug' denotes 'augmented', a number category which will be defined in section 2.2.1.

In Latvian (2), which continues the basic pattern of Indo-European nominal inflection (Benveniste 1935), nominal roots are extended into stems by the addition of a 'theme' vowel, and are inflected by the addition of a further affix, here *-m* 'dative.' This further suffix is variously referred to as *ending*, *desinence*, or *case-number* morpheme.[2]

Similarly in Huave (3), canonical transitive verbs and possessed nouns manifest an initial vowel, typically *a*, as well as Possessor Affixes (Pxs) or subject agreement affixes. Nouns in Latvian cannot occur without themes or desinences: **gulb* is not a well-formed word in Latvian. Likewise, Huave nouns are not well-formed without themes and Pxs: one cannot simply say **niing* 'house': one must say *a-niing-aran* 'somebody's house,' with an indefinite possessor suffix and theme vowel prefix.

The name applied here to such 'thematic' objects (the stem-final vowel in Latvian or the initial vowel in Huave) will be *stem formatives*. Stem formatives carry no syntactic information such as agreement or case features, nor do they affect a lexeme's projection into syntax. They are of a purely morphological character and comprise part of the well-formedness conditions on words as an idiosyncratic property of a given language.

Stem formatives may convey no morphosyntactic information at all. For example, Harris (1991, 1992) demonstrates that nouns and adjectives in Spanish require suffixes which he terms *word markers*. These markers are selected by the relevant root but express no morphosyntactic properties: they are word-pieces without function, yet required by autonomous word structure. For example, in (4) are illustrated examples of five of the seven such markers cited by Harris (1991:30-1):

(4)	a. *-o*	*muchach-o* 'boy (m.)'; *man-o* 'hand (f.)'
	b. *-a*	*muchach-a* 'girl (f.)'; *día* 'day (f.)'
	c. *-Vs*	*Luc-as* 'Luke (m.)'; *síntes-is* 'synthesis (f.)'
	d. *-u*	*espirít-u* 'spirit (m.)'; *trib-u* 'tribe (f.)'
	e. *-i*	*bikin-i* 'bikini (m.)'; *metrópo-li* 'metropolis (f.)'

As can easily be seen, the markers impose no gender: each marker may be part of a word of either gender. Furthermore, adjectival forms such as *mochal-es* 'loony, nutty,' and *curs-i* 'in bad taste' show that word markers do not carry category either.

[2]It has been suggested to me by Michael Kenstowicz that theme vowels may have arisen historically as strategies to assure that monosyllabic roots could be augmented so as to achieve minimal word weight, in the sense of McCarthy & Prince (1990a). Canonical roots in both Indo-European (Benveniste 1935) and in Proto-Huave (Suárez 1975) are monosyllabic.

Because stem formatives (or, word markers) exist, we are justified in assuming that there are conditions on the well-formedness of words which have nothing to do with the realization of inflectional categories. These well-formedness conditions comprise *autonomous morphological structure*. Autonomous morphological structure contains all of the conditions on word-level well-formedness which must be stipulated for a given language. This information resides in the Morphology Component, and is accessed in Phase II of the mapping from Syntax to Phonology. Autonomous morphological structure requires, for example, that Spanish nouns and adjectives must have contentless word-markers, that Latvian nouns have Themes and fused Case-Number desinences, and that Huave nouns have Themes and Possessor Affixes.

By hypothesis, stem formatives and inflectional affixes, while both figuring in autonomous morphological structure, are different sorts of word pieces distinguished crucially in the following way. While stem formatives may vary according to the properties expressed by inflectional affixes, they crucially express these properties only indirectly.

Consider the Huave forms in (5) and Latvian forms in (6):

(5) a. s - a - niing - an 'our (excl.) house'
1Px- TH -house- augPx

b. Ø - i - niing - an 'your (pl.) house'
2Px- TH- house-augPx

(6) a. zirg - a - m 'horse (dat.sg.)'
horse- TH -dat

b. zirg - u - Ø 'horse (acc.sg.)'
horse- TH- acc sg

In both Huave and Latvian, the theme vowel shows an allomorphy depending on other inflectional properties of the noun. In (5a) the theme vowel shows its default allomorph *a*, whereas in (3b), in the 2nd person, the theme vowel raises to *i*. In the Latvian forms, the theme vowel is default *a* in (6a) but raises to *u* in the accusative singular.[3]

[3]At first glance, it may appear that in *zirg-u-Ø*, the accusative singular desinence is actually *-u* and that the theme vowel *-a-* is deleted before a vowel, i.e. *zirg-a-u* --> *zirg-u*. But forms with theme vowel *-e-* also show raising in the accusative singular, e.g. *zem-e* 'earth,' *zem-i* 'earth (acc.sg.).' This shows that the final vowel in the accusative sg is really derived by a raising rule and not by its own (segmental) desinence. Cf. discussion in section 0.1.1.2.

The assumption is that stem formatives do not *directly* express the properties which condition their allomorphy, if they show any. But what 'directly express' means is a particularly intricate and theory-dependent question.

The claim that some morphosyntactic properties are realized directly through some phonological material and indirectly through others is an hypothesis of a particularly theoretical nature. Theories of inflection can be divided among those that (1) fully accept this hypothesis (Strict Lexicalism , Lieber 1980, Williams 1981, Selkirk 1982, Jensen & Stong-Jensen 1984, Jensen 1990; "No Autonomous Morphology" theories, Lieber 1992; and Weak Lexicalist Distributed Morphology, Marantz 1991, 1992, Halle 1989a, 1990, 1992); (2) partially admit it (Carstairs 1987); or (3) dispute it (Matthews 1972b, Anderson 1981, 1986, 1992).

There are a number of ways in which the distinction of direct vs. indirect expression of a property can be defined. We will consider this question with regard to the pair in (6). First, we consider one possibility within a Weak Lexicalist approach.

0.1.1.1 Halle's 1992 Model

In the analysis of Latvian nouns presented in Halle 1992, it is assumed that nouns must conform to a well-formedness condition illustrated below:

(7) [Noun + Q,Theme Vowel] + Q, Number-Case]

The Q represents an abstract dummy which occupies a string position until realized as phonological material by rules of Morphology known as Spell-Out. What is crucially assumed in this model is that morphosyntactic features of case and number are associated with the more external Q on the right, labelled Number-Case. An explanation of how this association comes about must be delayed until section 0.2. With regard to the pair *zirg-a-m* 'horse (dat.sg)' and *zirg-u* 'horse (acc.sg.),' the following assertions are valid:

(8)
a. The piece *m* replaces Q when Q is labeled Dative (Halle's 11c)
b. The piece *a* replaces Q, when Q is labeled Theme, for the class of lexemes including *zirg-* (8d)
c. Vowels become [+hi] before Q, Accusative-Singular (9a)
d. Q is replaced by Ø when labeled Accusative-Singular (11a)

In terms of properties and their realizations (in Carstairs' 1987 terms), the assertions in (8) indicate that (1) *m* realizes the property dative; (2) raising the theme vowel indirectly realizes accusative singular (inasmuch as these properties are part of the conditioning environment of the raising rule); and the absence of an ending directly realizes the properties accusative and singular. In (12) I give this formally.

(9) Direct vs. Indirect Realization

(1) If a (possibly null) string *S* replaces (spells-out) a Q linked to a property *p*, then *S* directly realizes *p*

(2) If *p* conditions a rule which does not spell out a Q to which *p* is linked, then the output of the rule indirectly realizes *p*

Certain assumptions of the model presented in Halle 1992 must be made clear from the outset. These are summarized below:

(10) (a) Each property *p* is linked to at most one Q

(b) Each Q occupies one and only one string position

The combined force of these assumptions is that each morphosyntactic property may have only one direct realization, analogous to what Carstairs (1987:151) terms its *principal exponent.*

Carstairs uses the following example from Latin to clarify the notion of principal exponence. In Latin the properties '2 sg' in the verbal inflection are usually realized as the suffix *-s* as in *ama:s* 'you sg love,' *ama:-bi-s* 'you sg will love.' But in the perfect tense (indicative), 2 sg is realized as *-isti:*, as in *ama:-v-isti:* 'you sg (have) loved.' Therefore, it might be argued that the suffix *-isti:* realizes all three properties '2 sg perfect.' Yet it is also the case that the perfect tense is signaled by the suffix *-v-*. According to Carstairs' definition, *-v-* will count as the principal exponent of perfect, since it unambiguously realizes perfect in some or all contexts where '2 sg' is not present. In other words, because there also exist such forms as *ama:-v-i* 'I (have) loved,' *ama:-v-it* '(s)he (has) loved,' etc. the affix *-v-* is independently established as the principal exponent of perfect and, although *-isti:* occurs only in the perfect, *-isti:* is only a secondary exponent of perfect, but a primary exponent of '2 sg.'

In this thesis, I will defend the view that each property has a (unique) principal exponent (in any given word), while debating a final assumption of Halle's (1992) model, given in (11):

(11) Each class of properties P is uniformly associated with a unique Q.[4]

This assertion is consistent with the Latvian nominal declension, for example. In the Latvian noun, Case and Number (classes of properties) merge together and so are uniformly associated with the rightmost Q, labelled Number-Case. Because each Q is associated with one and only one string position (10b), and each class of properties is associated with one Q (11), class of properties and position are inextricably linked in Halle's model.

However, while many languages admit a regular correspondence between class of properties directly realized and locations of realization, other languages require a more complicated analysis. In particular, in chapter 1 I argue that (11) should be assumed only as the default case. In other words, given no other rules of Morphology, a lexeme will regularly realize classes of properties in the same way, that is, it will display *system congruity* in terms of Wurzel 1989 [1984]. Latvian conforms to system congruity along this dimension since case and number are directly expressed as 'fused' suffixes without exception. Because system congruity is a tendency at best, probably associated ultimately with simpler grammars, deviations from system congruity are expressed here through additional or more complicated rules.

0.1.1.2 Strict Lexicalism: Affixes and Morpholexical Rules

In a Strict Lexicalist model, the distinction among directly realized and indirectly realized properties amounts to the difference between the features of an affix and the features of its subcategorization frame. Although many possible strict lexicalist analyses of Latvian might be proposed, the following is an approximate translation from Halle's (1992) analysis into an affix-based approach along the lines of Lieber (1980).

Since *m* realizes dative directly, and Ø realizes accusative singular, assume that the lexicon contains these entries:

(12) a. -*m* [+dative]
b. -Ø [+accusative singular]

[4]Note that this statement refers to classes of properties and not to particular properties. It should not be interpreted to mean that there cannot exist rules which change one property to another while holding the location constant, e.g. a rule which converts the genitive to nominative in the masculine animate noun in Russian.

Furthermore, in strict lexical theories such as that of Lieber (1980), stems are related to roots through morpholexical rule relations. Such a relation for Latvian *zirg-* would be:

(13) N ~ N -*a*

The morpholexical rule above relates noun roots to roots suffixed by *a*. (This would appear to be the default thematic vowel in Latvian nouns). Because the suffixes in (12) attach only to stems, they must subcategorize for stems, the suffixed alternants of (13):

(14) -*m* [+dat] [[-cons] ___]

Fortunately, thematized stems in Latvian are always vowel-final, so the subcategorization frame need merely specify that the preceding alternant end in a vowel ([-cons]).[5]

With respect to the raising of the theme vowel to *i* in the accusative singular, a number of options could be pursued.

The first option is to retain the zero suffix (12b) and to add a rule which raises the final vowel of any accusative singular noun. Hence *zirg-a* (acc.sg.) would input to the rule and surface as *zirg-u*.

The second option is to treat the raised theme as a stem alternant, by adding another morpholexical rule:

(15) N - [-cons] ~ N - [-cons, +high]

The zero-suffix of the accusative singular would then be required to subcategorize for the stem alternant with a [+high] vowel.

The third solution, reminiscent of Lieber's 1987 treatment of mutations in Fula and Welsh for example, would define the accusative singular ending as a floating [+high +round] autosegment:

(16) *[+high +round]* [+acc, +sg] , [[-cons] ____]

Upon suffixation of this autosegment, the final vowel would be raised, and if back, rounded. (Front vowels will not round because of Structure Preservation). It should be

---

5 Nevertheless, some account must be given for cases where the theme vowel deletes in the genitive and nominative singular for certain nouns classes (Halle 1992:41), as in *zirg-Ø-s* 'horse (nom. sg)' or *akmen-s* 'stone (gen.sg.)'

carefully borne in mind that such a change happens automatically only if it is assumed that affixation of autosegments entails that they link by tier conflation to the appropriate target, here the rightmost vowel, and that this linking can be expressed as a feature-filling rule. Because such linking is principle-driven, it need not be stipulated as a rule: all that need be known is where the floating affix is instantiated. In this way, floating (non-segmental affixes) do not differ from segmental affixes in any substantive way.

With the appropriate underspecification of vowels, it is possible to treat the accusative raising rule of Latvian as a feature-filling rule, as I will now illustrate. While it is not my aim here specifically to defend the radical underspecification required for this to work (see Archangeli 1984), the following summarizes the possible solution. Assume Latvian has the vowel inventory below, with the cooccurrence restriction 1 and the redundancy rules 2-6 at right:

	hi	back	round
a			
i	+	-	
e		-	
o			+
u	+		+

1. *[-bk, +round]
2. [ ] --> [+back]
3. [ ] --> [-round]
4. [+back, -round] --> [+low]
5. [ ] --> [-low]
6. [ ] --> [-hi]

If floating [+high +round] attaches to the vowels above before the application of the redundancy rule 3, then it will raise *a* to *u*. This follows because *a* is underlyingly unspecified for [±high], [±round] and [±low]. Once it becomes [+hi +round] through affixation, then it automatically escapes becoming unround by 3, low by rule 4, and nonhigh by 6. Floating [+hi +round] will also straightforwardly raise theme vowel *e* to *i* as in *zem-e* 'earth (nom.sg)' cf. *zem-i* 'earth (acc.sg)', before this vowel receives [-hi] by 6. Crucially, *e* will not round owing to the effects of Structure Preservation respecting feature cooccurrence restriction 1. Whether the array of features above is motivated by other aspects of Latvian phonology remains to be worked out, however, observe that it is crucial that the front vowels be underlyingly [-back] so that 1 may prevent rounding of *e*.

Underspecification is necessary only if affixation of floating autosegments must effect only feature-filling changes. If autosegmental affixes may have *structure-changing* (i.e., feature-changing) effects, then underspecification is partly unnecessary. We will return to this point in section 0.1.1.5. With regard to direct vs. indirect realization of properties, the lexicalist approach assumes that a given property, expressed by a feature or combination of features, is registered only once directly: an affix provides a feature value which percolates to the head of the word and all subsequent allomorphy refers to the feature, which after percolation becomes a property of the word as a whole.

To summarize, although the Weak and Strict Lexicalist approaches differ in means, they both distinguish principal and secondary expressions of properties.

0.1.1.3 Paradigm Theories: Rules without Affixes

Most paradigm theories dispute the distinction between principal and secondary properties. In the case of Latvian, a typical paradigm model such as Matthews 1972b would derive *zirg-a-m* and *zirg-u* by two disjunctive blocks of rules, ordered one after the other:

(17) I. a. Noun ---> Noun + *u* / gen.pl, instr.sg., acc.
b. Noun ---> Noun + *a* / elsewhere

II. c. Noun ---> Noun + *m* / dat.

The first block realizes what Halle's model labels theme vowels; the second, desinences. (This is of course only a fragment of the full set of realization rules necessary for Latvian).

Two things are especially noteworthy about the paradigm approach. First, each rule of affixation (or other process) is conditioned by features but does not in any way exhaust or discharge them. Thus, for example, the dative plural *zirg-ie-m* would require in this theory a rule of block I as below:[6]

(18) Noun --> Noun - *ie* / dat.pl, instr. pl.

The feature 'dative' employed in (18) must still be available to condition the rule suffixing -*m* in block II. Because no rule in this model exhausts or discharges a feature, there can be no distinction between principal and secondary exponents. The paradigm model easily permits the multiple realization of morphosyntactic properties, known as *multiple exponence.*

The second noteworthy characteristic of the paradigm model is that affixes never exist independently of the rules that introduce them. For example, the phonological string *u* introduced by rule (17a) has no status as an independent object of the theory. In other

[6]In a more complete analysis of Latvian, this rule would have to appear in a block after Block I, since -*ie*- may co-occur with a theme vowel as in *gulb-j-ie-m* 'swan (dat.pl)', where the theme vowel *i* becomes a glide *j* before -*ie*-. Cf. Halle 1992 fn. 9.

however, falls into a particular block of rules which, in order to be grouped disjunctively, have to be referred to. But *u* itself is nothing: no more or less privileged than the final *g* of the root *zirg*, for example.

In a paradigm model, it makes no sense to refer to the *-m* introduced on block II as the 'dative suffix' any more than one may call the *-ie-* introduced on block I (18) the 'dative plural' suffix. These assertions carry no weight because affixes neither carry features (as in the affix-based model), nor do they realize positions associated with features (as in Halle's model). Instead, they are introduced by rules conditioned by features.

0.1.1.4 Disjunctivity and Bleeding: The Discharge Effect

The three models entertained up till now each treat what I will call the Discharge Effect in a different way. Compare the two Latvian forms below:

(19) a. *zirg-a-m* 'horse (dat.sg)'
b. *ma:s-a-j* 'sister (dat.sg)'

The two nouns, both in the dative case, differ in that one has *-m* and the other *-j*. The latter affix is used with feminine nouns in the singular only, and the former affix *-m* is used elsewhere in the dative, including the plural, e.g. *zirg-ie-m* 'horse (dat.pl)', *ma:s-a:-m* 'sister (dat.pl.). All three theories accept and predict that one affix should be the 'default' realization of dative. Assume therefore that the default one is *-m*. The following three pairs of rules/representations express this idea in the three frameworks:

(20) Halle

a. Q --> *j* / [+fem] + ___, Sg.Dat.
b. Q --> *m* / ___ Dat.

(21) Affix-based models (including Strict Lexicalism)

a. *-j* [+sg, +dat], [+fem ____]
b. *-m* [+dat]

(22) Paradigm Model

a. Noun --> Noun + *j* / [+fem, +sg, +dat]
b. Noun --> Noun + *m* / [+dat]

In Halle's model, the first rule, (20a), is ordered before and bleeds (20b). It is ordered before (20b) because it is more specific. It bleeds (20b) because it rewrites Q by a string: (20b) cannot apply because after (20a) applies, the Q is gone. In this way, I will refer to these rules as *discharging* a Q, that is, rendering it inert to subsequent rules. For Halle's model, discharging a Q suffices to discharge the feature linked to it. I will show in this thesis that Qs (positions) and features may be discharged independently, and so Halle's model must be appropriately modified to permit this eventuality.

In the affix-based model, discharge is effected more indirectly. The affixes *-j* and *-m* cannot both affix to a single noun stem either because (1) they both belong to the same disjunctive lexical stratum (or, nearly equivalently, the same syntactic bar level); or (2) the same feature cannot be percolated twice. The first option is straightforward and entirely analogous to the disjunctive blocks of the paradigm model. The second is trickier. Lieber (1990, 1992) assumes, for example, that lexical classes have associated with them feature structures called "categorial signatures." When an affix such as *-j* affixes to a stem, the features supplied by the affix percolate to the word as a whole and lodge in the categorial signature. Assume therefore that nouns in Latvian have at least the signature as in (23):

(23) ___ sg
___ dat

The blanks are feature coefficients which are to be supplied by affixes. When *-j* affixes to *ma:sa-*, the features [+dat] and [+sg] percolate to the signature and fill the slots in the signature. Thereafter, no affix may attach which supplies a feature already supplied. Subsequent affixation of *-m* is impossible. We can then speak of percolation as *discharging* a slot in a categorial signature.

Affix-based models must make an additional assumption, however. There is no reason as yet why **ma:s-a-m* could not be derived as the dative singular form, with *-m* [+dat] attaching and then [+sg] being filled in by default at then end of the derivation. Lieber (1992:101), for example, assumes that the 'default (negative) values of the remaining unspecified features are filled in' after all affixes have attached (cf. also the Unmarking Principle of Jensen 1990:141).

Some principle must be invoked to ensure that the most specific affix possible is chosen. Jensen (1990) invokes a version of the Elsewhere Condition to derive this effect:

(24) Elsewhere Condition (Jensen 1990)

> If two or more morphemes can be attached to a third form, the morpheme with the more specific subcategorization frame takes precedence and blocks the attachment of the other morpheme. A subcategorization frame, A, is more specific than another, B, if A specifies everything that B specifies and more.

However, it is important to note that this principle is completely at variance with representation-based approach. The Elsewhere Condition, as originally formulated by Kiparsky (1973), was a principle of *rule* application. Jensen's Elsewhere Condition requires a global purview of all potential affixes and a choice among these. But the guiding insight of the representation-based model is that affixes attach *freely* to compose words, subject only to their own subcategorization requirements, and not to the subcategorization frames of other affixes. Appealing to the Elsewhere Condition in the lexical model requires that an affix may attach to a stem only if another more specific affix *could not* have attached.

This problem becomes even more acute in Lieber's more radical proposal (1992), in which all word-formation takes place in the syntax and there is no autonomous morphological component. It makes no sense to speak of an Elsewhere Condition like Jensen's operating over sets of X^0s in syntax. This would formally entail that the only well-formed sentences would be those projected from the maximally selective lexical items. Intransitive predicates, for example, requiring one argument, could never occur, since transitives, having a more specific selectional requirement, would always 'win out.' Obviously this leads to absurdity.

One way to avoid this problem here would be to subcategorize *-m* for the complement of environments for which *-j* subcategorizes. But then it is impossible to retain the idea that *-m* is the default dative affix. Because any affix-based model must include a hierarchy among affixes, but to do so must appeal to a principle like the Elsewhere Condition, entailing a global purview of affixes, its conceptual foundation is considerably weakened. In this thesis, I will defend a theory in which morphosyntactic representations are discharged by means of affixes, essentially giving a lexicalist model in reverse. However, because the representation is the input to the mapping, it suffices to say that the most specific affix matching the representation is chosen in cases where there is competition. Since rules are required to instantiate affixes on this model, rules can be order-d according to complexity and a hierarchy of features (to be introduced), giving the required affix hierarchies.

The Paradigm Model enforces discharge by means of the Elsewhere Principle as a condition on rule disjunctivity (Kiparsky 1973). The rules in (22) are disjunctively ordered, meaning that only one will apply to any given form, the most specific rule possible. As we will see later, the paradigm model crucially links discharge to the rule block, which is typically, though not always, confined to providing exponence at a particular position. Each rule block may have one and only one final default rule. I will show in this thesis that the result of these two ideas is that the discharge effect for any given feature (i.e. the domain in which it may condition a rule only once) is limited to the same domain as that of some single final default rule. These domains are not in fact coextensive and I therefore will abandon the paradigm model as strictly interpreted.

In particular, we will see that to duplicate the effects of discharge across blocks of rules, Anderson (1986, 1992) must have recourse to identity rules, i.e. rules which rewrite a string as itself (cf. also remarks in Halle 1990, Marantz 1992). Such unnatural contingencies are unnecessary in the theory advanced here. To avoid such problems, Anderson (1986), as I will show, introduces disjunctivity across rule blocks for certain types of rules. This is done by means of a particular definition of disjunctive application. The theory to be presented here captures the effect of disjunctions across rule blocks directly without appeal to such a specialized definition.

To summarize, the three models so far exhibited enforce discharge effects through different means. Halle invokes bleeding of Qs; Affix-based models fill categorial signatures and/or block less specific derivations by global principles; Paradigm theories have recourse to the Elsewhere Principle. More specifically, Halle discharges Qs (and, by implication, the features linked to them), while Lieber (1980, 1992) may discharge positions by means of strata/bar-levels, and features by means of filling categorial signatures. Paradigm theories discharge only positions (inasmuch as positions are correlated with rule blocks), leaving features intact, and, where features must be discharged separate from positions, identity rules are invoked.

In this thesis I will be focusing a great deal on the notion of discharge. I will be arguing that positions and features may both be discharged (rendered inert to future rules). A discharged position is filled by a string and cannot be filled again. (This duplicates Halle's Q-rewriting, and can be understood as filling a template position). When a feature is discharged, its principal exponent is realized. Thereafter, the feature may condition allomorphy rules, but cannot again have a principal exponent. (This duplicates percolation). The full theory I have in mind will only become clear after chapter 1 in which I analyze the Afroasiatic prefix-conjugation in some detail.

0.1.1.5 Defining the Rule-Affix Distinction

In this section I will present a theory-internal definition of the rule-affix distinction. We have seen in the previous sections that existing theories of morphology tend to be generally either affix-based or rule-based. Representation-based or affix-based models such as Lieber (1980,1992) permit the expression of processes such as ablaut and mutation as the attachment of floating affixes which fill in features left underspecified in stems. In this way, many rules can be reduced to affixes. The paradigm-model, on the contrary, excludes the notion affix entirely, treating each affix as a rule introducing the string of the affix.

It should be carefully noted that the rule-affix distinction is really a matter of structure-building vs. structure-changing effects. An affix -- whether segmental or floating -- in merely *adding* material to a pre-existent form, will always have a structure-building effect. A rule, on the other hand, may have a structure-changing or feature-changing effect.

If floating affixes are allowed to have feature-changing effects, that is to say, the features they carry can *overwrite* features already represented on the unaffixed form, then the distinction between what counts as a rule and what counts as an affix is almost entirely lost. To see this, consider again the Latvian accusative raising rule presented in 0.1.1.2. Recall that the theme vowels /a,e/ raise to /u,i/ in the accusative singular. I presented an underspecification of the Latvian vowels that would allow this raising rule to be effected by a floating autosegment [+high +round] 'accusative', which, without changing any features, fills in unspecified features of the theme. But consider now what possibilities may arise if the affix has feature-changing effects.

The final *a* of *zirg-a* can be fully specified as [+low, -hi, +back, -round], whereupon a floating [+hi -low +round] accusative singular affix can in a *feature-changing* fashion convert *a* to *u*. Observe, however, that the change to from *e* to *i* must respect the inventory of Latvian and not create a rounded vowel *ü*. (The rule then would be structure-changing, i.e. feature-changing, while the effect would be structure-preserving, i.e. inventory-preserving. For the distinction of these two uses of the term 'structure preservation,' see Mohanan 1991).

Such feature-changing affixes are akin to "melodic overwriting" affixes recently defended, for example, by Yip (1991). It is of paramount importance, however, to observe first, that on such a view, autosegmental (floating) affixes must be designated as either feature-filling or feature-changing, and second, that there is now little if any distinction between rules and representations. Any rule operating at a morph boundary can be reinterpreted as a feature-changing floating affix, which automatically creates the same effect as a rule (with the exception of metathesis rules).

Consider for example the alternation *shelf, shelves*. The stem modification voicing the final fricative can be seen as a case of 'overwriting association,' if it is assumed that for a limited class of nouns, the allomorph of plural *z* selected has a feature [+voice] ([+slack]) which may overwrite the [-voice] ([+stiff]) specification of the final *f* in *shelf*. Non-alternating nouns such as *puff, puffs* would not select the feature-changing "overwriting" affix.

In this thesis I will defend the idea that the distinction between feature-changing rules and structure-building affixes should be maintained. Feature-changing rules (overwriting affixes) do not count as principal exponents of morphosyntactic properties, which I hypothesize are confined to structure-building operations (filling positions with strings, adding features or prosodic material). Feature-changing rules (overwriting affixes) are always secondary exponents, which are expressed by dynamic rules.

These ideas form the following working hypothesis, which I adopt:

(25) The Rule-Affix Hypothesis

(1) A morphological rule introducing an affix is a structure-building rule. Structure-building rules discharge features and positions-of-exponence. The affix so introduced is the principal exponent of the features discharged.

(2) A structure-changing (or feature-changing) rule does not discharge any features or positions-of-exponence.

To summarize, the autosegmental alternative most successfully blurs the distinction between affixes and rules. Whether this is a welcome result is certainly open to question. I will choose to regard the feature-changing vs. structure-building distinction as the criterion for dividing rules from affixes. The consequence of this is that readjustment rules and morphophonemic rules of various types do not count as "affixes," and are dynamic. Not being objects but rather processes, they cannot carry or discharge features, nor do strings affected by them discharge positions-of-exponence.

For example, German umlaut accompanies certain plural forms, e.g. *Haus, Häus-er*, 'house(s).' On the assumption that the umlaut rule is a feature-changing rule, it cannot discharge (or "spell-out") [plural]. The realization of an overt affix, whether segmental or autosegmental, does discharge [plural], as long as it is feature-filling only. One consequence is that we should not expect to find cases where a feature-changing rule bleeds an affix.

The most notorious alleged counterexample to this claim is the English noun plural. It would appear that ablauts such as *foot - feet, mouse - mice*, bleed regular plural affixation in *-s*. However, because other nouns fail to take *-s* in the plural, but also fail to ablaut, e.g. *sheep*, it is not necessarily ablaut which blocks plural in *-s*. The nouns which have no overt plural affix discharge plural by means of -Ø and this bleeds *-s* ; on the current analysis, ablaut does not bleed *-s*, since it is a feature-changing rule and therefore cannot discharge plural. The analogous situation in the strong (ablauting) verbs is even more illustrative:

(26)

	past tense		past participle		
	ablaut	suffix	ablaut	suffix	
	+	+	+	+	*los-t; sol-d; tol-d*
	+	-	+	+	*spoke, -en; froze, -en*
	+	-	+	-	*won; dug; wound (< wind) ; shone*
	+	-	-	+	*ate, eat-en; saw, see-n, blew, blow-n*
	+	-	-	-	*came, come; ran, run*
	-	+	+	+	*swell-ed, swoll-en; shear-ed, shor-n*
	-	+	-	+	*show-ed, show-n; lade-d, lade-n,* reg.verbs
	-	-	-	-	*cut, hurt, hit, cost, set, shed, shut*
	-	-	-	+	*beat, beat-en*
*	-	-	+		
*	+	+	-		
*		+		-	

As can be seen above, ablauting cooccurs with overt suffixes in both the past tense and in the past participle of verb forms. Two generalizations emerge from this distribution. First, if the past tense is neither ablauting nor overtly suffixing, the past participle is not ablauting; inversely, if the past tense is both ablauting and suffixing, so is the participle. Second, if the past tense is overtly suffixing, so is the past participle. These generalizations govern the distribution of suffixation and ablauting in the past tense vis-à-vis the participle, but, as far as a given past or participle form is concerned, ablaut and suffixation are freely cross-classifying. Thus, the gaps in the above array have nothing to do with the bleeding of a suffix by a rule of ablaut.

A similar example from Ket, a Yeniseyan language spoken in central Siberia, will demonstrate that the English case is not an accident. Ket nouns form plurals in a variety of ways (Dul'zon 1968:68-69):

(27) a. -ŋ (50%)
b. -n (43%)
c. Vowel ablaut: ses', sas' 'river(s): река, -и'
d. Vowel ablaut and ending: tet, tat-n 'husband(s): муж, мужья'
e. Stress shift: qaiŋóq , qáiŋoq 'fly, flies: муха, -и'
f. Tone shift: dès, dê?s 'eye(s): глаз,-а'

Most Ket nouns -- about 94% -- form the plural either with a velar or coronal nasal suffix, and this is partly predictable from gender and phonology. But a small percentage -- about 7% -- form plurals by ablaut or prosodic adjustment.

Because the tonal and stress properties of Ket remain poorly understood, it is impossible to determine whether the plurals in (27e,f) are effected by feature-filling autosegments (e.g. floating tones). What is important to note is that, as in English and German, plurals formed by ablaut can have no overt ending (27c), or have an overt ending, as in (27d). The ablaut rule would seem to bleed affixation in the cases of (27c), but not in the cases of (27d). On my account, inasmuch as ablaut is a feature-changing rule in Ket, then ablaut cannot discharge plural. It is thus to be expected that some ablauted forms will also have an overt plural ending, as they indeed do: tat-n 'husbands.'

The Ket plurals are in this way rather similar to the variety of strong, weak, and doubly-marked plurals of English and German. This variety is not an accident of Germanic.

Kurylowicz's first law of analogy is relevant here. It can be summarized: "a bipartite morpheme tends to replace a unitary one" (Arlotto 1972:137). The example of a bipartite morpheme used to illustrate this law is the German *Gäst-e* 'guests,' where the plural form has both an ablaut and an affix, cf. singular *Gast*. An etymological unitary-marking, as in OHG. *Topf-e*, analogizes to the doubly marked form, giving Mod. Germ. *Töpf-e* 'pots,' cf. *Topf* 'pot.' However, it is significant here that the doubly-marked form has a *rule* of ablaut along with an *affix* as the principal exponent of plural. Kurylowicz's law follows automatically for such cases once it is understood that the analogy is merely the extension of the ablaut rule to cover more plurals than before. On my analysis, 'plural' is not in such cases a bipartite morpheme; rather, the ablaut rule accompanies plural as a secondary exponent. What is not expected is for there to be multiple affixes of a single property, since one affix will discharge the property, making it impossible for another affix to realize that property directly.[7]

Stem-adjusting vs. suffixing plurals are widespread in Semitic, most notoriously in Arabic. Arabic nouns form plurals either by syllable-structure adjustment ('broken plurals') (McCarthy & Prince 1990b) or by the addition of overt suffixes ('sound plurals'); adjectives may have both a broken and a sound plural, whose distribution is determined by the morphology and semantics of the noun it modifies (Thackston 1984:19):

(28) Broken Plurals

a. nafs --> nufūs 'soul'
b. ?asad --> ?usūd 'lion'
c. karīm-at --> karā?im 'noble'
d. fākih-at --> fawākih 'fruit'
e. kitab --> kutub 'book'
f. rukb-at --> rukab 'knee'

(29) Sound Plural

muxliṣ --> muxliṣ-ūna (m.pl.nom.); muxliṣ-īna (m.pl. obl.) Sound
muxliṣ-āt-un (f.pl. nom.); muxliṣ-āt-in (f. pl. obl.)

(30) Broken and Sound Plural (Adjective)

a. kabīr --> kibār 'big' Broken
b. kabīr --> kabīr-ūna (m.pl. nom.); kabīr-īna (m.pl. obl.) Sound
kabīr-āt-un (f.pl.nom); kabīr-āt-in (f.pl. obl.)

What is significant about the Arabic case is that syllabic readjustment as in (28) is never accompanied by overt suffixation, *nor* are there plural forms in which there is neither syllabic readjustment nor overt suffixation. A noun either has a broken plural or a sound plural, but never both (at once) or neither. In terms of acquisition, then, the child will not be motivated to assume that the broken plurals have a contentful -Ø suffix marking plural, since this suffix, unlike English -Ø 'plural' (e.g. *sheep, moose*) would never exist apart

[7]The so-called 'double plural' phenomenon of Breton, Welsh, and the diminutives of Yiddish are apparent counterexamples. Yet these are really instances of the same affixes figuring as number desinences functioning as stem augments. For a discussion of this phenomenon with attention to 'perfective' Latin verbal stems, see Aronoff (in progress, ch. 2). Cf. also the 'plurals of abundance' in classical Arabic, which are broken plurals formed from broken plurals, e.g. *kalb* 'dog,' *?aklub* 'dogs' *?akālib* 'dogs (pl. of abundance).' (Wright 1971:232, McCarthy & Prince 1990b:220).

from the syllabic readjustment rule. Under such circumstances, one might be inclined to suppose that the readjustment rule does indeed bleed overt suffixation. However, there is an alternative analysis.

In cases where a structure-changing rule is invariably disjunctive with an overt affix, as in the Arabic broken plurals, I propose that the feature discharged by the affix is in fact an *inherent* feature of the stem in the instances without overt suffixation. Thus, the Arabic broken and sound plurals have the following structures:[8]

(31) a. [N, pl]
b. [N] [pl]

The sound plurals have the feature 'pl' outside the N constituent, whereas the broken plurals have the feature 'pl' inherent to them. To have a feature inherently means that the feature is supplied to the form before syntax. Inherent features arise either as listed elements on a given form (e.g. gender of nouns in many languages), or by rules of the lexicon (such as redundancy rules supplying gender or inflectional class to forms by various criteria). In Chapter 2 I will show that number is inherent to many nouns in the Kiowa-Tanoan languages. In Arabic, nouns which participate in the broken plural alternation are freely supplied with inherent plural number lexically; this inherent feature conditions the appropriate prosodic form of the broken plural by later rules. An inherent plural specification suffices (in Arabic) to block the further application of [pl] in syntax.

Observe also that in the sound plural cases, (31b), case may or may not be fused with plural: in the masculine sound plurals, case is fused with plural, while in the feminines it is separately realized. In the inherent, broken plurals, case cannot fused with number because number is inherent to the noun.

If the [pl] feature of broken plurals is in fact inherent then it is predicted that it should somehow behave differently than a [pl] feature which is supplied syntactically (perhaps through a Number head in the sense of Carstens 1991) and is realized as an overt ending. This is indeed the case as can be verified from the so called 'deflected' agreement pattern of broken plurals.

The agreement patterns of Arabic plurals nouns are complex and have varied from Koranic to modern times and in different registers (Thackston 1984: 17ff, Haywood & Nahmad 1965:43 ff.). The general pattern for Koranic Arabic is that nouns referring to human plurals agree according to their natural gender, and nouns referring to thing plurals

[8] I would like to thank Alec Marantz for this suggestion.

agree according to their grammatical gender if sound plurals and *as feminine singulars* if broken plurals. We will concern ourselves only with the grammatical agreement for nouns referring to things. Because the sound masculine plural is used only of nouns referring to male humans, there are but three types of thing plurals:

(32) Agreement of Nouns Referring to Things in Koranic Arabic

gender	pl type	agreement	example
masc	broken	f sg	kutub-un kabīr-at-un 'books big-f.sg'
fem	sound	f pl	jann-āt-un maʕrūš-āt-un 'gardens trellised-f.pl'
fem	broken	f sg	mudun-un kabīr-at-un 'cities big-f.sg'

Observe that the broken plurals referring to things agree as feminine singulars, whereas, at least in Koranic Arabic, sound (feminine) plurals referring to things agree normally as feminine plurals. (This usage is declining in favor of treating all feminine plurals, including those referring to women, as singulars, but this is a separate and later development). The agreement pattern shown above extends to verbal agreement as well (although for independent reasons verbs agree for number only when preceded by their subjects).

On the basis of these agreement facts, it may be concluded that the Arabic broken plural is not merely one of two possible ways of morphologically signalling the property plural, but rather acts as a special type of agreement class, associated with feminine singular agreement. This lends support to the idea that broken plurals are in fact inherently plural, that is, their [pl] feature is supplied in the lexicon to the stem itself, in the same component that assigns default gender and inflectional class. I then proposed that as an inherent feature, an inherent [pl] (like, for example, the feature [mass]) may block the assignment of [pl] from syntax, thereby preventing forms which have plurals both by affixation and prosodic readjustment.

Masoretic Hebrew provides a striking contrast to classical Arabic in this area. In Hebrew, syllabic readjustment is not disjunctive with overt suffixation, as shown below (Lambdin 1971:9, 17):

(33) Suffixation Alone

a. bəḵôr --> bəḵôr-îm 'first-born'
b. ḫălôm --> ḫălôm-ôṯ 'dream'

(34) Suffixation with Readjustment

a. mélek̲ --> məlāḵ-îm 'king'
b. mizbēᵃḫ --> mizbəḫ-ôṯ 'altar'

In (33), the two plural suffixes occur without syllabic readjustment or ablaut; in (34), both syllabic readjustment/ablaut and overt suffixation occur. Under such circumstances, the child will not be prompted to conclude that 'pl' is inherent to the stem as in Arabic, but rather than a rule of stem readjustment accompanies the realization of plural in a certain class of nouns.

It should be borne in mind that a rule freely supplying inherent 'pl' to the stem must be stipulated in Arabic, and therefore comes at come cost. Only when positive evidence is presented will the child acquire this rule. This evidence is the absolute disjunctivity of syllabic readjustment and overt suffixation in Arabic and the peculiar 'deflected' agreement pattern of broken plurals. Hebrew shows neither of these phenomena and therefore does not have inherent number.

To summarize, rules *change* phonological structure or phonological features, but cannot discharge morphological features or positions. Thus, a rule will never realize a morphological feature's primary exponent. *Affixes* on the other hand, discharge morphological features, making them invisible to further rules except as conditions for allomorphy. The distinction between the two types rests on the more primitive distinction between information-adding vs. information-changing processes. Rules may appear to function as a feature's primary exponent, as in the case of the Arabic broken plural, but this is an illusion. I proposed that in such cases, the feature 'pl' is in fact inherent to the stem, as evidenced by the fact that broken plurals behave as a special agreement class.

0.2 Morphosyntactic Well-formedness

The theory assumed here takes the input to (inflectional) Morphology to be the output of syntactic rules. Two operations occur in Syntax which are crucial for deriving the inflectional shape of words: (1) X^0 movement and (2) Concord. The former process, familiar from the work of Travis (1984), Baker (1985), creates complex syntactic objects by the adjunction operation. X^0 Movement also raises V to INFL (or an expansion of INFL into several heads), whereupon the agreement and tense features of V are either provided or checked, depending upon the theory espoused, Emonds (1976), Chomsky (1989), Pollock (1989). The second process, Concord, copies morphosyntactic features from one syntactic projection to another, as in adjective-noun concord.

Additionally, two rather significant adjustments take place to the structures provided by syntax. The first is the rebracketing of X^0 projections under adjacency. These rebracketings/mergers are discussed by, among others, Sproat (1985) and Marantz (1988). I will also assume that syntactic structures are not linearized, since linearization is a property of phonological form only. Linearization and Merger apply simultaneously, respecting adjacency requirements of X^0s. Merger is the process whereby an X^0 may exchange its requirement of adjacency to a constituent with the merger relation with the head of this constituent (Marantz 1988). What counts as a 'head' depends upon the level at which the merger operation takes place; within syntax 'head' has its usual meaning, so that Merger defines head-to-head movement; after syntax, the head of a constituent may become any peripheral element.

The second adjustment is Impoverishment (Bonet 1991), in which fully-specified morphosyntactic representations are simplified before they input to Morphology. Because Impoverishment deletes certain features, it effects the systematic neutralization of categories. I discuss Rebracketing, Merger, and Impoverishment in the following two sections.

0.2.1 Syntactic X^0s and the Positions of Exponence of their Features

The relation of X^0s in syntax and of the ever elusive morphological object "morpheme" has been variously interpreted, usually with implicit adjustment in the definition of the latter term. It will be useful to disambiguate and specify these views before going any further.

Syntacticians tend to interpret "morpheme" to mean any X^0 head which combines into a word with another head by X^0 movement. For example, Giorgi & Pianesi (1991) give the following structure for a complex verb of Classical Latin:

(35) [[[[[*lauda-* $_V$] *-v-* $_{T2}$] *-e-* $_V$] *-r-* $_{T1}$] *-am* $_{AGR}$]
laudaveram 'I had praised'

For Giorgi & Pianesi, 'morpheme' is coextensive with syntactic head. Each subpiece of the complex verb has its own projection in syntax, and the entire verb is combined by means of successive X^0 movements.[9] The phonological form of these pieces is usually ignored in such work, and complex questions of allomorphy are not addressed at all. As Carstairs-McCarthy has put it (1992:152), "[f]or the time being, allomorphy remains for them [proponents of syntactic affixation] something of an embarrassment." One exception to this general trend is the work of Lumsden (1987, 1992). This thesis deals directly with this neglected area, while retaining wherever possible the advances obtained in the theory of syntactic affixation.

Aside from allomorphy, a more serious problem is that the relation of X^0s to word-pieces is not always direct. Consider the examples from Classical Arabic in (36):[10]

(36)
a. *t-aktub-aani* 'they (f dual) write'
f- write-dual

b. *n-aktub-u* 'we write'
1 pl-write-IMPF

c. *y-aktub-na* 'they (f pl) write'
DFL -write-f pl

As I will show in detail in section 1, the prefix and suffix positions of the Arabic imperfect can be filled by affixes whose feature content is quite varied. For example, the prefix position is filled by gender in (36a), but by person and number in (36b), and by a default prefix in (36c). The suffix position is filled by number in (36a), by tense in (36b), and by gender and number together in (36c). Morphologists tend to use the term "morpheme" to

9The movement of the V through a functional head T and then to a lexical head V here is in violation of a constraint on X^0 movement discussed in Li (1990), unless T and AGR are somehow 'lexical' in Latin. However, this is a possible assumption on a certain interpretation of lexical, given that T discharges a thematic role (cf. Speas 1990). I leave this topic open.

10DFL stands for the 'default' prefix, which fills a position of exponence just in case no other more specific affix does.

refer to affixes (sometimes structure-changing rules) such as *n-* '1 pl' or *-na* 'f pl.' A more precise term for such word pieces is *morph*.

I will avoid the term *morph* since it has a particular definition rooted in American Structuralism, whereby each morpheme is a set of morphs or actual realizations of the morpheme. I will use a more neutral terminology, referring to word pieces such as *n-* '1 pl' simply as *affixes*.

It is important to observe that the affixes in (36) (as principal exponents) express both AGR features and TENSE features. To explain this sort of distribution, I will be assuming that as part of Rebracketing/Merger, syntactic heads may fuse into single objects, all of whose features may contribute to the realization of affixes. This will be termed the Fusion operation, which takes two merged syntactic heads and yields a single terminal position.

If AGR and TENSE (formerly, INFL) are separate projections in syntax, as is assumed by Chomsky (1989), and for Latin, by Giorgi & Pianesi (1991), the output of V raising to AGR and TENSE is held to be:

(37) [[V TENSE $_{T}$] AGR $_{AGR}$]

Assuming suffixal Linearization, this configuration may become (38), where [*] represents strict adjacency:

(38) [[V] * TENSE] * AGR $_{AGR}$]

At this point, the head AGR can exchange its adjacency relation to the constituent to its left with the merger relation with the head of this constituent:

(39) [[V] * TENSE +AGR]

This process is formally identical to that of cliticization, where rebracketing and merger has grossly altered the syntactic structure of the string (adapted from Marantz 1988:253):

(40) a. [$_{IP}$ I [$_V$ will [$_V$ go [$_{PP}$ to Milwaukee]]]]

$\downarrow$

[[I'll] [go [to Milwaukee]].

b. [$_{DP}$ le porc-épic [$_{PP}$ de [$_{DP}$ le garçon]]]

$\downarrow$

[le porc-épic [[du] garçon]]]
'The boy's porcupine'

For the purposes of Morphology, rebracketed constituents such as [T+AGR] in (38) or the constituents instantiated by *I'll* or *du* in (40a,b) comprise *fused* constituents. As such they display behavior as if they were atomic: differentiation of tense features and agreement features is no longer enforced in [T + AGR], and *I'll* and *du* behave as single domains (morphological words). For expository purposes, I will refer to the fused T +AGR unit as INFL.

There exists no agreed name for such objects, aside from the overused term morpheme. For absolute clarity, I find it necessary to abandon the term morpheme, since it is when undefined unbearably vague, and even when defined explicitly, prone to endless misinterpretation. The unlabelled constituents as in (38) I will call morphosyntactic constituents, or M^0s. The reader is advised to take careful note of this. As used in this thesis, M^0 is not the same as X^0 except in the default instance, nor is it the same as position of exponence, nor is it the same as an affix (strings, autosegmental tones, etc.) or set of affixes which occupy a position.

0.2.1.1 Splitting of M^0s

The Arabic case is then complicated by the fact that the M^0 INFL *splits* phonologically into two positions of exponence: prefix and suffix. The diagram below schematizes this process:

(41) [V [T + AGR $_{M^0}$]] unlinearized

____ verb stem ____
prefix suffix

The crucial point of the Arabic example is that X^0s and positions of exponence (the prefix and suffix positions) are not isomorphic. It is misleading to conflate them. While some languages display a certain isomorphy or near-isomorphy between what may be postulated as syntactic heads and positions of exponence (so-called "agglutinative" typology), this need not be the case. In this thesis, I will be developing a theory which pays strict attention to the degrees of divergence of syntactic objects (X^0s) and the positions of (principal) exponence for the features carried by X^0s. It will be my contention that isomorphy between X^0s and positions of exponence is the default situation, complications of which must be expressed by additional stipulations of grammar, most typically constraints on morphological well-formedness as described in section 0.1.1.

For example, I will be arguing that in Arabic, the prefix and suffix positions are demanded by the V+INFL constituent as part of its *autonomous word structure*. These positions are required solely by conditions of *morphological* well-formedness. The positions must be filled by phonological material (including contentful -Ø as in the jussive mood), in order for the constituent to become a well-formed word. The positions are the analogues to the Qs of Halle's model, described in the previous section. Crucially, however, in the present theory Qs are isomorphic to X^0s only in the default case. Additionally, as I will show, the prefix and suffix positions are separate elements: the Arabic tense/agreement "morpheme" is not simply a circumfix or simulfix that splits into two pieces, encircling the stem. Rather, the prefix and suffix are realized separately and interactively.

A more complete analysis of Arabic (and many other Afroasiatic languages) will be undertaken in chapter 1. In chapter 3 I show that autonomous word structure may be of three sorts: (1) strictly licensing, as in Arabic, where specific positions in excess of the number of M^0s are required for well-formedness; (2) unmarked, where there is isomorphy between M^0s and stems/affixes; and (3) freely licensing, where M^0s are 'liberated,' such that they split into indefinitely many non-obligatory positions.

Of course, it may seem entirely redundant that T and AGR are assembled by Merger/Fusion, and then split again into two affixes. Why not indeed postulate a single syntactic projection INFL? Since the primary focus of this thesis is not on the arguments for or against expanded INFL, I merely note here that both Merger/Fusion and Splitting are permitted on the current proposal. Expanded INFL is convenient and possibly correct, insofar as the syntax of expanded INFL projections is isomorphic to the syntax of the phonological pieces which are the (principal) exponents of these projections, as argued by Giorgi & Pianesi for Latin (and Italian). Insofar as INFL remains atomic morphologically (fused, or fused and then variously split again, as in Arabic), a hypothesized expanded

INFL must either be assembled by Merger in morphology, or not postulated in the first place. For now I leave this question open.

0.2.2 Capturing Systematic Neutralizations

It is a quite obvious fact that some languages morphologically express more morphosyntactic properties than others. To give a simple example, verbs in Semitic express the gender of the subject in some person-number categories, while this is typically not the case in Indo-European. Similarly, it is easily observed that whatever morphosyntactic properties are available for a category such as Verb rarely cross-classify fully with each other. For example, gender in classical Arabic is never distinguished for 1st persons, nor for 2nd person dual categories. The paradigm below illustrates these neutralizations:

(42) The Arabic Prefix Conjugation (Imperfect Indicative) *k t b* 'write'

singular	dual	plural	
y-aktub-u	y-aktub-aani	y-aktub-uuna	3 m
t-aktub-u	t-aktub-aani	y-aktub-na	3 f
t-aktub-u	t-aktub-aani	t-aktub-uuna	2 m
t-aktub-iina	*	t-aktub-na	2 f
ʔ-aktub-u	*	n-aktub-u	1

(I will be presenting the paradigms with 3rd person at the top, in the fashion traditional to Semitic studies).

In (43) I list some facts about the paradigm above:

(43)
a. There is no gender distinction in the 1st person
b. There is no dual number in the 1st person
c. There is no gender distinction in the 2nd person dual
d. There is no 1st person inclusive vs. exclusive
e. The 3rd fem and 2nd masc are the same in the singular and dual
f. Gender is masculine or feminine but never neuter

While each of these facts holds of the paradigm, it would be incorrect to attribute all of them to the same grammatical mechanism. In what follows, I will make explicit the different ways in which these generalizations are encoded.

A number of issues are at play here. We can take care of the generalization in (43e) immediately, it results from a homophony of the prefixes *t-* '2' and *t-* 'fem.' Aside from this accidental homophony, the other syncretisms are more systematic.

0.2.2.1 The Feature Alphabet and the Category Alphabet

Each language must possess a set of morphosyntactic features such as person, number, and class features of various kinds, including sex-based gender and grammatical gender-class (including such properties as physical shape, deixis, animacy, whether real or classificatory, and even phonological shape (Aronoff 1992)). Just as the phonology of a language picks out certain of the set of universal phonological features to be active in defining its lexical alphabet of segments, so too must a language pick out a set of morphosyntactic features. Some morphological features, such as features for conjugational class, may be language-specific, however.

How is this selection to be accomplished? Or, put more appropriately, how is such a selection learned by a child?

I will be following in large part certain assumptions made in work by Calabrese (1988, 1992) on the composition of phonological alphabets. Here the ideas he presents will be adapted with some modification to apply to what I will term a "morphological alphabet."

Calabrese's view is that Universal Grammar provides lists of features and feature complexes arranged in hierarchies according to their complexity. The hierarchies are actually defined as lists of feature cooccurrence restrictions, henceforth *filters*. When a child is provided with an auditory stimulus that a certain feature combination is phonologically active in a language, the filter which bars this combination is suppressed.

Acquisition evidence suggests that very young children acquire the phonological systems of their language by exposure to the clustering of sounds at certain acoustical positions. Research beginning with Eimas et. al. (1971) and Streeter (1976) and fully developed in experiments carried out by Werker (1989), has revealed that infants perceive a highly articulated inventory of phonological distinctions, but, in response to speech in their environment, between the ages of 6 to 12 months, they lose the ability to perceive certain distinctions. For example, a child exposed to only English will lose the ability to distinguish retroflex coronals from alveodental coronals and aspirated from non-aspirated voiceless stops.

Acoustical clustering presumably provides the cue that certain distinctions are operative. For example, if the stimulus language makes a contrast between aspirated and non-aspirated stops, then the child will not hear sounds clustering at the border between the two, that is with voice-onset timing highly variable, but rather will hear clustering away from the acoustic boundary. This provides the clue that aspiration is distinctive within the sound system.

The youngest infants perceive a multitude of distinctions, that is, they are ready to perceive whatever distinctions may happen to be present in the speech environment. I hypothesize, therefore, that at different stages of very early development, (approximately 6-8 months) the filters postulated by Calabrese (1988, 1992) appear and are 'suppressed' in the event that exposure to the relevant distinction has already occurred.

It is predicted, then, that the filter barring [+spread] from appearing with [+cons, -cont] appears in infants at approximately 6-8 months and is suppressed in the infants that hear Hindi and not suppressed in the infants hearing English only.

A particularly surprising result of Werker (1989) concerns children who heard Hindi spoken in early childhood, whose first words were Hindi, but never went on to acquire Hindi fully but instead learned English as adults and remembered little or no Hindi in adult life. This group of subjects was found to perceive the aspirated/unaspirated distinction nearly as well as Hindi-speaking adults. This confirms the developmental hypothesis. The suggested conclusion is that a filter is presented once in infancy and if suppressed at that time, it is never 'unsuppressed' even if the filter is operative for other speakers in the adult-language speech community.

In a filter-based theory, certain implicational relations can be expressed as follows. It has been observed that no language distinguishes [ATR] for high vowels unless it distinguishes [ATR] for mid vowels. Similarly, no language distinguishes [ATR] for low vowels unless it also distinguishes [ATR] for nonlow vowels. Calabrese (1992:14) proposes the following (partial hierarchy):

(44) *[-ATR -high]
|
*[-ATR +high]
|
*[+ATR, +low]

If the child's stimulus is sufficient to suppress a filter, all filters above this filter are suppressed automatically as a consequence of UG. If a child perceives that [ATR] is contrastive for low vowels, then ATR must be contrastive for all heights. It follows that the burden of learning consists solely in finding that point on the hierarchy which is the complexity threshold, that is, the lowest filter which must be suppressed.

Converting these ideas into the realm of morphosyntax, I will assume that there exists a universal set of features for at least person, number, tense and aspect. Gender features (Corbett 1991) and the classification of verbal action called Aktionsart are possibly universally defined but show such variability and are so poorly understood that I will have little to say about these here.

Of the universal set of morphosyntactic features, some are active in a given morphosyntactic system, while others are inactive. Of the active features, not all will fully cross-classify with each other, giving the effect of 'gaps' in inflectional paradigms. Filters account for such gaps.

Some of the facts cited in (43) for Arabic follow from the selection of active features while others follow from filters. (43e) -- that Arabic verbs do not have an explicit neuter gender -- follows from the choice of active features which combine lexically to form arguments and agreement heads in syntax.

On the other hand, since gender and dual number cross-classify with some but not all of the other categories, generalizations (43a-c) are better expressed as filters:

(45)	a.	*[1 f]	No 1st person feminine
	b.	*[1 dual]	No 1st person dual
	c.	*[2 dual f]	No 2nd person feminine dual
	d.	*[1 2]	No 1st person inclusive

I propose here that filters of this sort, like phonological filters, are part of UG and, at a certain developmental point, will be acquired automatically unless evidence is provided which contradicts them.

All four filters in (45) are active in familiar contemporary Indo-European languages, which have no gender in (most) verb forms, no dual, and no inclusive.

Consider now the 1st person feminine. This category must be both 1st person and feminine at the level of agreement, since an adjective agreeing with a feminine 1st person will appear as feminine:

(46) ʔana mustaʕidd-at-un
I prepared (fem.)
'I (f.) am prepared' (Abboud & McCarus 1985:291)

A more familiar example would be the French alternation *Je suis heureux/heureuse* 'I am happy (m./f.).'

Now given the filter in (45a), one should expect that the features [1] and [f] should never be allowed to combine. The facts of agreement however, show that syntactic structures are more fully specified than the differentiation of morphological forms would betray at first glance (cf. also Lumsden 1987, 1992).

To express this situation, I will assume that, of the features selected as syntactically active in a language, some are deleted at the input to Morphology. This process will be called Impoverishment, following Bonet (1991). I will be arguing in the next several subsections that Impoverishment occurs automatically in the first phase of Morphology whenever combinations of features violate filters such as those in (45). I will propose that a hierarchy of features determines which feature(s) of the offending combination is deleted to assure well-formedness at Morphology.

0.2.2.2 Feature Hierarchies and Feature Trees

In the morphological realization of 1st person feminine V+INFL in Arabic, the deleted feature is [f] and not [1]. We know this because 1st person feminine V+INFL is not homophonous with 3rd person feminine V+INFL, as might be expected if [1] were deleted. (I will assume that 3rd person lacks a person feature specification in Arabic: details later). This deletion could be expressed by a rule as below:

(47) f --> Ø / 1

However, the rule mechanism is too powerful to express this sort of systematic ambiguity, since it could just as easily express the opposite relation:

(48) 1 --> Ø / f

Rule (48) would incorrectly render 1st person feminines homophonous with 3rd person feminines in Arabic.

Instead of rules of this type, I propose that the features in a filter are not simply unordered but actually are composed in a hierarchical structure. If 1 > f in this hierarchy, then by a principle, when the filter *[1 f] is activated (at Morphology), f and not 1 will be deleted.

Moreover, if such hierarchies are represented graphically as tree structures, then the deletion of morphosyntactic features can be assimilated to Delinking of autosegmental representations, as first proposed by Bonet (1991):

(49)
AGR
|
1
|
f

For Bonet, Impoverishment is limited to Delinking, and the geometries she proposes are motivated to allow the permitted impoverishments and no others. For example, using this same idea, 1 in (49) cannot be delinked unless f is also delinked. In accordance with the filter *[1 f], it follows that the minimal modification necessary to comply with the filter is delinking f only.

It is not obvious however, that the hierarchies which tell you which feature is deleted by Impoverishment are deducible from an invariant morphosyntactic feature geometry.

First consider an example from Calabrese (1992:5 ff.). Calabrese notes that the Italian pronunciations of German *führer* [fü·rər] vary between [firer], [furer], and [fiurer]. The former two cases are examples of delinking one offending feature in the filter *[+round -back]. If [+round] is deleted, the the result is [i]; where [-back] is delinked, the result is [u]. (The third case represents breaking, discussed shortly). Following the Sagey (1986) model of feature geometry, Calabrese assumes the model of articulator-based feature geometry as below:

(50)
PL
LAB — DORS
[+round] — [-back]

Since neither feature dominates the other, the restriction that filter violations must be repaired by delinking forces neither the choice of delinking [+round] nor the choice of delinking [-back].

We can conclude that while feature geometries in phonology express what sorts of delinkings are possible, feature geometries do not express which delinkings are necessary.

Consider all of the filters in (45) and their putative geometries. If in each case the hierarchy which determines which feature is to be deleted is expressed by the relation of dependency in a tree, we are led to postulate these structures:

(51)

a. AGR	b. AGR	c. AGR	d. AGR
│	│	│	│
1	1	2	1
╪	│	│	╪
f	pl	pl	2
	╪	│	
	dual	dual	
		╪	
		f	

(I am assuming for now that dual is a subclass of plural, since 1st person dual neutralizes with 1st person plural, not singular). In each case, the hierarchy below holds:

(52) $1 > 2 > \text{pl} > \text{dual} > \text{f}$

Hierarchies of features such as the one above have been proposed before. For example, Silverstein (1976) proposes a similar hierarchy to explain the various types of split ergativity which may occur. Lumsden (1987, 1992) has also argued that hierarchies of features are necessary in any theory in which 'the forms of inflection are underspecified.' In particular, Lumsden argues that the feature [±pl] must be higher than [±fem] in Romanian and suggests that this ordering 'might be predicted on universal grounds' (p. 1992: 480). In this thesis, I propose that feature hierarchies of the sort suggested by Silverstein and Lumsden can be used to explain the direction of neutralization by filters, as well as the unmarked (least costly) order of morphological rules (chapter 1) and of affixes (chapter 3).

The idea that morphosyntactic features are not simply unordered lists is not at all new. In particular, Gazdar, Klein, Pullum & Sag (1985) (and references included there) experiment with morphosyntactic feature geometries within the theory of Generalized Phrase Structure Grammar. At this point, however, it remains unclear what purpose a morphosyntactic feature geometry should have aside from expressing hierarchies as in (52). As we have seen, the hierarchies that motivate delinking in phonology are independent of the dependency relation in phonological feature geometry,

It is possible to identify three competing criteria underlying arguments for feature geometries. First, and earliest, is Clements' (1985) proposal that only subtrees operate as natural classes of features for phonological rules. The node dominating the relevant subtree defines this natural class for autosegmental spreading. The second criterion is anatomical: the nodes dependent on PLACE in particular are associated in Sagey's (1986) model with specific articulators of the vocal tract; features associated with each such node (LABIAL, CORONAL, DORSAL) are solely executed by this articulator. Finally, the relation of dependency automatically encodes the property of contrastiveness. For example, the feature [±distributed] depends on CORONAL in the feature tree; therefore, [±distributed] can appear (be contrastive) only for trees which have the node CORONAL from which [±distributed] may hang. Crucially, however, not all matters of contrastiveness are encoded by the feature geometry. For example, the fact cited by Calabrese that [±ATR] is contrastive for high vowels only if it is contrastive for mid vowels is nowhere expressed by the geometry.

In the realm of morphosyntactic features, the first of these criteria has a translation. (At present I see no analogy to 'morphological articulators'). Gazdar, Klein, Pullum & Sag employ this first criteria in constructing feature-geometries for morphosyntactic features (1985:21): 'whole clusters of feature specifications [can] be picked out in a natural way if they share[d] a mother node in the graph.' The idea that only subtrees operate as natural classes corresponds to the idea that (1) only subtrees may operate together as features of an affix and (2) only subtrees may be copied in Concord operations. (Only the latter usage is cited by Gazdar et. al.)

The first consequence makes wrong predictions immediately. Consider the Arabic verbs in (53) and the putative geometries of their AGRs in (54):

(53)
a. t-aktub-na — 'you (f, pl) write'
2-write-pl f
b. t-aktub-aani — 'they (f, du) write'
f-write-dual

(54)
a. AGR | 2 | pl | f
b. AGR | pl | dual | f

In (53a) the prefix expresses person while the suffix expresses number and gender. If the prefix *t-* is limited to being the principal exponent of some subtree of (54a), then it should be the principal exponent not merely of 2 but also of all the features which 2 dominates. Yet this is clearly false, since the suffix *-na* expresses f pl separately. Similarly in (53b), f, although dominated by dual, is realized as a separate affix.

On the basis of such examples, we must assume either that subtrees are not the natural classes of features which affixes may realize or that the putative geometries are incorrect.[11]

As for the second consequence-- that only subtrees may be copied in concord relations -- this remains to be investigated and is more promising. For example, agreement with adjectives (in Arabic) copies only the subtree dominated by number, bringing any gender with it. I leave this topic open.[12]

The idea that the tree should express contrastiveness is ultimately linked to the idea that some phonological features are articulator-bound and therefore contrastive only when dependent upon that articulator. No obvious analogy with morphosyntax exists and complications arise immediately when this analogy is forced. For example, the hierarchy 1 > 2 predicts that 1st person inclusives merge with 1st person exclusives whenever there is a systematic ambiguity of these categories.[13] Nevertheless, if the same dependency relation is held to express contrastiveness, then 2 should appear only when dependent on 1, or, put another way, 2 cannot appear without 1. Therefore, a simple 2nd person category should be impossible.[14]

To summarize so far, morphosyntactic feature trees express the hierarchies which determine which feature in a filter is deleted by Impoverishment, as summarized by the claim in (55) and the principle in (56):

(55) *Systematic Neutralization*
Systematic Neutralizations are the result of Feature Cooccurrence Restrictions (filters). Filters combined with feature hierarchies effect Impoverishment.

---

[11]However, as far as I am aware, it is possible to maintain the idea that only *subgraphs* of these trees figure in the realization of affixes. Specifically, in the Arabic imperfect, no affix is the principal exponent of person and gender to the exclusion of an intervening number feature. This leads the possibility that morphological realization of affixes consists in tracing a (contiguous) path through a graph of morphosyntactic features.

[12]Cf. Greenberg's Universals 32 and 36: "Whenever the verb agrees with a nominal subject or nominal object in gender, it also agrees in number; ... If a language has the category of gender it always has the category of number."

[13]Algonquian languages show a syncretism of 1st inclusive and 2nd person for some Pxs and verbal prefixes, but this is a separate issue which I address in section 2.1.5.

[14]Note that the person features I am assuming are privative. I do this because the value [-1], at least, is not active morphologically, as pointed out first (in entirely different terms) by Zwicky (1977). A complete discussion of morphosyntactic features of person and number will be given in Chapter 2.

(56) *Impoverishment*
Given a filter of the form *[aF bG], delink that feature which is lower on the hierarchy of features.

Any theory of morphosyntactic feature geometry is in its infancy at this point. Whether there is in fact a convergence of the criteria of natural classes for rules and of hierarchy for neutralization is far from settled. For the purposes of this thesis, I will be assuming that hierarchies exist among morphosyntactic features, but remain uncommitted as to whether geometric representation of this hierarchy is in any way useful or valid.

0.2.2.2.1 Filters are not Constraints on the Content of Affixes

The filters necessary for conditioning the neutralization of categories are not simply constraints on the content of possible affixes, but rather are constraints on the features borne by M^0s as a whole.

To see this, consider the constraint *[1 dual] in Arabic. Let us suppose momentarily that this constraint does not hold of the INFL complex as a whole but rather holds only of possible affixes. In other words, there can be no affix bearing both the features [1 dual]. It is a mistake to suppose that this interpretation of filters ensures the correct neutralizations, as I now show.

Whenever an M^0 splits into two positions of exponence, it is possible that the features barred by a filter may be realized separately by well-formed affixes. For example, compare the Arabic forms in (57):

(57)
a. ʔ-aktub-u 'I write'
1-write-IMPF

b. t-aktub-u 'you (m) write'
2-write-IMPF

c. t-aktub-aani 'you (dual) write'
2-write-dual

d. *ʔ-aktub-aani 'we (dual) write'
1-write-dual

Although the affixes ʔ- '1' and *-aani* 'dual' are both well-formed, they cannot co-occur to form a 1st person dual form as in (57d). It is not sufficient, then, that *[1 dual] hold of affixes. Instead, the filters which express systematic neutralizations must hold of entire M^0s.

The position taken here differs somewhat from that of Lumsden (1987, 1992) as I now explain. Lumsden (1992:470ff) carefully distinguishes 'positions versus morphological signals.' He argues that the morphological signals, that is (loosely speaking), the affixes themselves, are necessarily underspecified in many cases. For example, if the German pronoun *was* 'what' is underspecified for case features, this explains why it may appear in a case-conflict environment where explicitly case-marked *wer* 'who (nom.)' and *wen* 'whom' cannot. In the following examples, cited by Lumsden (1992) from Taraldsen (1981), the matrix verb *zerstöre* 'destroy' requires an accusative direct object, but the subject of the free relative clause must have nominative case:

(58) Ich zerstöre was/*wer/*wen mich ärgert.
I destroy what/*who/*whom me annoys

'I destroy what/who annoys me.'

Lumsden concludes that 'the morphological signal is always nondistinct from the specifications of the underlying position, but it does not always express every distinction that is specified in that position.' This is entirely consistent with both the EWP theory and with Halle's abstract morphology: realization or spell-out rules do not detail exhaustively the context in which they apply: instead, their structural descriptions are *minimal*, giving the effect of 'underspecification' of morphological signals.

However, Lumsden's theory does not include the notion of Impoverishment. All neutralizations are expressed by the underspecification of morphological signals. In a theory in which (syntactic) positions, i.e. X^0s, and the affixes (morphological signals) which realize these positions are isomorphic, this might be possible. But once splitting of the Arabic type is introduced, Impoverishment must be introduced as well. There is no way to neutralize [1 dual] into [1 pl] in Arabic by means of underspecifying the morphological signals which realize an underlying [1 dual] form, since as I have shown, fully underspecified ʔ- [1] and -aani [dual] exist independently in the system and should be expected to combine. I therefore conclude that not only may morphological signals be underspecified (as seems now uncontroversial), but also that the inputs to Morphology undergo simplifications which may express highly systematic neutralizations. These Impoverishments apply irrespective of the inventory of affixes (or realization rules) of a language.

Impoverishment filters do not have feature-changing effects; rather, they simply remove syntactically operative distinctions from the purview of morphological rules. I address the question of feature-changing readjustments in 1.17.

If both Impoverishment and the 'underspecification' of morphological signals exist, then both might logically be available to express a given neutralization. It is expected, then, that one sort of neutralization type might lead to the other through historical change. For example, consider the following schematic paradigm of case endings:

(59)

	declension 1		declension 2	
	sg	plur	sg	plur
nom.	a	d	g	j
gen.	b	e	h	k
acc.	c	f	i	j

In this hypothetical language, nom pl and acc pl are identical in declension 2 [j] but distinct in declension 1 [d,f]. The ending [j] may accordingly be underspecified, being a morphological signal for at most plural and 'direct' case for declension 2. Now in a theory in which there is no Impoverishment, an analogical change whereby nom pl comes to be identical to acc pl in declension 1 as well as declension 2 can be expressed only as the 'loss' of the [f] ending from the set of affixes/realization rules. But such a theory predicts that any of the declension 1 endings [a,b,c,d,e,f] are equally plausible candidates for loss from the system. Analogy is therefore never predicted.

On the other hand, a theory which includes Impoverishment makes a different claim. Suppose first that the neutralization of nom and acc in declension 2 is reanalyzed as an Impoverishment, namely *[declension 2, pl, ACC], entailing the deletion of ACC. It is then possible for this filter to be extended to declension 1 as well: *[ACC pl]. This extension is natural, since it is a simplification of the filter. The spread of nom-acc syncretism can then be explained as the simplification of the filter. The 'loss' of ending [f] follows automatically.

Attributing systematic neutralizations to filters predicts not only that filters may change but also that they may be gained or lost through historical change without significant reanalysis of the affixes themselves or of well-formedness conditions on word shapes. This prediction is in fact true as I will show in chapter 1. In particular, I show that the analogue of ill-formed (57d) is well-formed in the South Arabian languages Meḥri and Soqoṯri, which have lost the *[1 dual] filter.

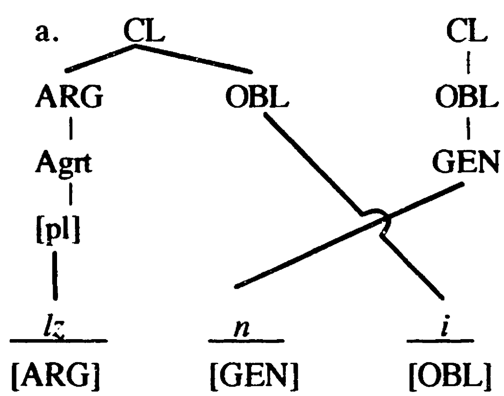
0.2.2.3 Breaking and Negation

In Calabrese's theory of phonological complexity, there exist three means for resolving filter violations. (Such violations arise when a filter is added to a grammar through historical change, when a rule produces a segment which violates a filter, and in the pronunciation of foreign segments). These are: (1) Delinking, (2) Breaking, and (3) Negation. I have discussed Delinking already.

Breaking splits a complex segment into a contour segment whose parts are simpler, permitted segments, such as the breaking of German [ü] into [iu] by speakers of Italian. Finally, Negation switches the values of both features in the filter, for example lowering [+hi -ATR] [ɩ] to [-hi +ATR] [e] in early Italian, or in the change from [ü ö] [-back +round] to [+back -round] [ɨ ə] in Mongolian (Calabrese 1992:29).

It is unclear whether Breaking and Negation are necessary in morphosyntactic Impoverishment, although Delinking is clearly necessary. Bonet (1991) analyzes the fission of certain clitics in Catalan as examples of breaking a clitic (CL) tree into two parts. To give one example, Bonet argues that the 3rd plural oblique clitic, realized as *lzi*, can break into its 3rd plural part *lz* and its oblique part *i*. That Breaking has indeed taken place is shown by the fact that in between the two parts may appear another clitic, genitive (partitive) *n* :

(60) a.



(= Bonet 1991:111, (36a))

b. *de pomes, als nens, no* [lzə ni] *donis!*
of apples, to-the children, 3rd,pl+part.+dat., give

'do not give apples to the children!' (= Bonet 1991:111, (36b))

This fission, however, is not a product of complexity per se, but rather the result of linearization of the phonological exponents of the clitics according to principles specific to this dialect of Catalan, Barceloní. As such, I treat such examples as cases of splitting of M^0s, like the splitting of Arabic INFL into prefix and subject positions.

However, it is clear that phonological breaking rules are also sensitive to constraints on proper sequencing of the broken parts. For example, consider the following examples of breaking rules from Pre-Old English short vowels to Old English short diphthongs (Cassidy & Ringler 1971:30-1):

(61) a. Before liquid + consonant or before [h],
*æ > ea
*e > eo (if liquid is [l], consonant must be [k] or [x])
*i > io

b. Examples
*hærd > heard 'hard'
*herte > heorte 'heart'
*tihhian > tiohhian 'arrange'

The breaking rules shown above do not indiscriminately linearize the components of the broken segment: in particular, the [+back] segment of the resulting diphthong is always the second element. Calabrese (1992:26) has hypothesized that whenever breaking results from complexity, the first half of the broken diphthong is the primary feature. I discuss primary vs. secondary features in the next section.

Phonological breaking and M^0 splitting share the properties of taking a complex and atomic unit and dividing its content between two separate locations in a string, dependent upon constraints on the sequencing elements. Since Breaking rules in phonology are well-attested, an analogous phenomenon in the morphosyntax-morphology mapping strengthens the claim made here that morphosyntactic and phonological features have similar rule systems.

I have found no examples where Negation plays a role in categorial neutralization. However, a number of morphological rules discussed in chapter 2 rely crucially on feature-switching or α–notation operations. These include the rules for the distribution of enclitic -*a* in Mam (2.1.6), a Mayan language, and the Kiowa-Tanoan "number-switch" rules (2.3).[15]

[15] Also to be included here may be the so-called 'chiastic' concord of numbers in Semitic, whereby masculine nouns take feminine number modifiers and feminine nouns take masculine number modifiers.

0.2.2.4 Enhancement and Feature Transparency at Morphology

Calabrese incorporates the idea from Stevens & Keyser (1989) that in certain feature combinations, one feature is primary while the other feature value enhances the first. For example in the segments [u,o], the feature [+back] is primary while the feature [+round] phonetically enhances the backness of the vowel.

The morphological analogue to this can be seen in the number features of Arabic. While in most cases dual is a subclass of nonsingulars ("plurals"), there are examples in which duals are formed from the singular stem and not the plural stem. Consider the examples below from the Arabic perfect (suffix-conjugation):

(62)	a. katab-at	'she wrote'
	b. katab-at-aa	'they (f dual) wrote'
	c. katab-na	'they (f pl) wrote'
	d. katab-ti	'you (f) wrote'
	e. katab-tum-aa	'you (dual) wrote'
	f. katab-tum	'you (pl) wrote'

The 3rd person feminine dual (62d) is formed from the singular verb *katab-at* 'she wrote,' whereas the 2nd person dual form (62e) is formed from the plural verb *katab-tum*. Therefore, we can expect the 3rd and 2nd person duals to have the representations in (63):

(63)	a. [dual]	3rd person
	b. [-sg, dual]	2nd person

The question then becomes how to ensure that [-singular] is absent in the 3rd person but present in the 2nd person. Consider the status of the matrix [-sg dual]. At LF, where number features are interpreted, the value [-sg] is otiose in the presence of [dual], since two-membered sets are by implication non-singleton. Similarly, in a language in which all nonlow back vowels are round, i.e. one in which the filter *[+back -round] / [+low] is not switched off, the feature [+round] is predictable, and serves merely to phonetically enhance the backness of the vowel. Redundant values of this sort are typically transparent for phonological rules such as vowel harmony (Steriade 1987). In Arabic, the plurality of the verb in the second person form *katab-tum-aa* is an enhancement of the duality of the form. Such an enhancement need not be present, as (62b) shows.

In terms of geometry, such cases are problematic. If dual depends on plural (as it must by the hierarchy, since 1st person duals neutralize with plurals, not singulars), then it is predicted that dual cannot appear if [-sg] is not there in the tree. Thus (63a) cannot be derived from (63b), since if [-sg] is delinked, then so will [dual] which depends on it. We may say that in geometric terms, dual is *uniquely licensed* by [-sg], just as in phonological geometry for example, [+distributed] is uniquely licensed by CORONAL:

(64)
[-sg]	COR
|	|
[dual]	[+distributed]

Wherever dual is present, so must [-singular], so dual will always be enhanced by plural, contrary to fact. In other words, instead of *katab-at-aa* 'they (f dual) wrote', we should expect **katab-na-yaa* or some other similar form (and in fact such forms do occur in Ugaritic, as will be shown in section 1.14). Of course one might suppose that *-na* 'f pl' is specifically [-dual], but this flouts the generalization that [-dual] is never a natural class morphologically.

Because of such examples, I will propose that features may be absent from morphological representations in two possible ways. The first is via Impoverishment, which obeys the hierarchy. Impoverishment always deletes non-redundant values. The second is to never allow certain redundant values into the representation in the first place, so they cannot be present at morphology (and so never condition affixes or rules). Such will be the case with the feature [-sg] in *katab-at-aa*. Enhancing (logically predictable) features may be absent, but need not be.

The Arabic case can be handled with the following filter:

(65) *[-sg dual] / 3rd person

The filter in (65) requires that in the 3rd person, dual arguments cannot be enhanced by the redundant specification [-sg]. This filter will be active in the lexicon and not merely at morphology. In other words, no 3rd person AGR can have both [-sg] and [dual], although having only [dual] or only [-sg] is possible (giving unenhanced dual and plural, respectively).

An analogy to a similar case in phonology will illustrate how plural can be said to enhance dual in some circumstances, but not others. Consider the filter *[+sg dual]. Because no argument can be simultaneously [+sg] and [dual], since this is logically incoherent, this feature combination is absolutely impossible, just as is *[+hi +low] in phonology. Therefore, the value [-sg] is redundant in the presence of [dual] and can be filled in automatically as an enhancing feature. This enhancing feature may have a morphological effect, e.g. the 2 plural suffix *-tum* in addition to the dual suffix *-aa:* in (62e). Thus we may say that the duality of the form becomes more perceptually salient since it is reinforced by [-sg] in *-tum*.

If, however, a language has an additional filter *[-sg dual], then neither value of [±sg] can be filled in. Thus the dual must remain unspecified as to [±sg] all the way through to LF. There is no particular *semantic* effect of this underspecification, since [-sg] is fully otiose semantically whenever [dual] is present.

Similarly, a language may possess the filter *[+round], that is to say that rounding will not be distinctive at all. In such a case, it will suppress the filter *[-round +back] / [+low], and the back vowels will all be redundantly [-round], that is, not enhanced by rounding. This would be analogous to a language in which duals were never enhanced by [-sg] and thus were always formed from the stem form unmarked for number. (Such cases are indeed rare, as are languages which permit no round vowels.)

In Huave, only the mid back vowel /o/ is round, and all other vowels are [-round]. The inventory of simple vowels is: /i e ɨ o a/. This can be derived by assuming that Huave has the filter *[-round +back] / [-hi -low], which guarantees that the mid back vowel will be enhanced by [+round]. *[+round +hi] guarantees that /ɨ/ will be unround. The Huave vowel /o/ is like the 2nd person dual in Arabic: its enhancing feature is permitted to surface. The Huave vowel /ɨ/ is like the Arabic 3rd person dual: the enhancement it is expected to have by UG is not present, owing to a special, language specific filter. The filter *[+round +hi] in Huave corresponds to *[-sg dual] / 3rd person in Arabic.

In sum, certain redundant values may be omitted from morphosyntactic representations by language-specific filters. I illustrated this by showing that logically predictable [-sg] is absent in the 3rd person dual in Arabic. I argued that these redundant values are absent from morphosyntactic representations in virtue of filters which hold throughout a derivation. Thus, [-sg] is not permitted in the presence of [dual] (in the 3rd person) at any time in the derivation, and so [-sg] and [dual] never combine at all. Thus, [-sg] need not be erased by Impoverishment. Because of the hierarchy of filters, it would be expected that if [-sg] were deleted, then its subsidiary feature [dual] would also be deleted. This does not happen since in fact [-sg] is not deleted but rather is never part of the representation at all in the 3rd person.

0.3 Summary

There exist conditions on the well-formedness of inflected words which are independent of syntax and phonology, as is evidenced, for example, by the theme affixes of the archaic Indo-European nominal inflection and of Huave verbs and nouns, by the obligatory fusion of case and number into a single desinential position in Indo-European, and by the word-markers of Spanish. I took these as the starting point for a theory of autonomous morphological structure, which is provided in a module of grammar between the output of syntax and the input to phonology.

In this module called Morphology, rebracketing (merger) and the simplification of morphosyntactic representations called Impoverishment, essentially "prepare" strings of syntactic heads and their features for phonological realization as well-formed words. I argued that features are discharged ("spelled-out") in virtue of conditioning the filling of positions-of-exponence with affixes, but I showed that in the Arabic prefix-conjugation, features of a given type (deriving from the same syntactic source) need not appear uniformly in the same position of exponence. Indeed, the relation between minimal syntactic projections (X^0s) and positions-of-exponence (Qs) is not one-to-one except in the default instance. I argued that theories such as Lumsden (1987, 1992), in which sets of affixes are in competition for realizing syntactic positions, or Halle (1989a, 1992), where syntactic heads are spelled as strings, must be modified to permit the possibility of 'morpheme splitting.'

Thus, features, positions and affixes are the three objects manipulated by morphological rules. Traditional theories of morphology combine all three into the *morpheme* -- an entity consisting of a particular type of features at a particular position realized as a set of affixes -- and in so doing fail to recognize the intrinsic independence of features, positions and affixes. The theory to be presented in this thesis explicitly recognizes these three objects and their independence, and thus differs significantly from all other theories.

In addition, I argued that the difference between rules and affixes should be reduced to the difference between feature-changing and feature-adding processes in phonology. When morphological processes such as ablaut and metathesis change features or have structure-changing effects, then they are rules and do not discharge features of the input representation. When a morphological process has only a feature adding effect, it may discharge features of the input, i.e. "spell" these features as phonological material. Such material need not be segmental: that is to say, floating autosegments may be affixes so long as their effect is feature-filling only.

I also presented a theory of morphological feature neutralization whereby input morphosyntactic representations are Impoverished in the sense of Bonet (1991). Impoverishment deletes the lower feature(s) in a combination of features violating a filter at Morphology. What counts as "lower" depends on a universal hierarchy of features. These filters determine the complexity ceiling for a given category type in a given language. I argued that these filters are analogous to the filters in Calabrese's (1988, 1992) theory of phonological alphabets. Morphological filters, however, create the morphological alphabet: the set of categories which are realized inflectionally. I proposed, following Calabrese, that filters are part of Universal Grammar and that, at a certain point in maturation, children suppress them when exposed to (positive) evidence that the filters are violated. Such a theory allows a straightforward interpretation of the cross-linguistic markedness of inflectional categories.

In the next chapter I provide a detailed analysis of the prefix-conjugation in Afroasiatic, including much discussion of the historical changes that this highly archaic inflectional paradigm has undergone. In this context, I will elaborate the theory sketched above and compare it to its competitors.

CHAPTER 1:
The Afroasiatic Prefix Conjugation and Discontinuous Bleeding

1.0 Introduction

In this chapter I investigate some key issues beginning with the classical Arabic prefix-conjugation as the basis of discussion. In sections 1.1 - 1.2 the prefix-conjugation is introduced and a naive preliminary analysis of its affixes is given.

In the following sections (1.3-1.7), I analyze the Arabic imperfect according to three prominent models of inflection: Anderson's Extended Word-Paradigm theory (1981, 1986, 1992, and much other work), the Strict Lexicalist position, specifically that of Lieber (1980, 1989, 1992), and Halle's abstract morpheme model (1989a,1990, 1992). By contrasting these three models, we find that, when viewed at the appropriate level of abstraction, each has recourse to many of the same devices to express generalizations such as positional disjunctivity, blocking, syncretism, and empty affixes.

I identify three entities which play a role in constructing word-forms: morphosyntactic features, phonological material (affixes), and, crucially, positions of exponence. While all theories recognize the former two as primitives which figure in the word-formation rules of the theory, each ties the last-- position-- to a single device: disjunctive rule block (Anderson), stratum (Lieber), syntactic head X^0 (Halle). In doing so, each theory fails to directly capture what I call "discontinuous bleeding," where Elsewhere Principle effects occur *across* position-classes or rule blocks.

To capture discontinuous bleeding, I propose that morphological words may select autonomous morphological structures which include *obligatory* prefix/suffix positions. These positions are identified with the Qs of Halle's model, which serve as "dummies" occupying template positions until these are identified with syntactic material. However, the Qs of the model presented here differ crucially from Halle's Qs insofar as they are not isomorphic to syntactic heads. Inflectional morphology results from rules that place features from syntax into Qs and then either (1) delimit the Q prosodically, i.e. *build* structure into it , or (2) rewrite the Q as a string. The former rule type gives the effect of morpheme-structure constraints, including truncation, and the latter type gives affixes.

In sections 1.12 through 1.17 I present comparative and historical evidence for the firmness of discontinuous bleeding relations. Finally, in section 1.18, I analyze the prefix-conjugation of Berber in the proposed theory, contrasting this with an analysis in the Extended Word-Paradigm model. The historical extension of the affix -ɣ in three Tuareg dialects is shown to follow automatically on an analysis which relies crucially on discontinuous bleeding.

1.1 The Prefix Conjugation Paradigm

The Classical Arabic prefix-conjugation is illustrated below.

(1) Imperfect

singular	dual	plural	
y-aktub-u	y-aktub-aani	y-aktub-uuna	3 m
t-aktub-u	t-aktub-aani	y-aktub-na	3 f
t-aktub-u	t-aktub-aani	t-aktub-uuna	2 m
t-aktub-iina	*	t-aktub-na	2 f
?-aktub-u	*	n-aktub-u	1

The forms above are given for the imperfect aspect. The subjunctive and jussive forms are shown in (2) and (3):

(2) Subjunctive

singular	dual	plural	
y-aktub-a	y-aktub-aa	y-aktub-uu	3 m
t-aktub-a	t-aktub-aa	y-aktub-na	3 f
t-aktub-a	t-aktub-aa	t-aktub-uu	2 m
t-aktub-ii	*	t-aktub-na	2 f
?-aktub-a	*	n-aktub-a	1

(3) Jussive

singular	dual	plural	
y-aktub	y-aktub-aa	y-aktub-uu	3 m
t-aktub	t-aktub-aa	y-aktub-na	3 f
t-aktub	t-aktub-aa	t-aktub-uu	2 m
t-aktub-ii	*	t-aktub-na	2 f
?-aktub	*	n-aktub	1

The prefixes are invariant with respect to tense/mood, but the suffixes alternate as follows. Wherever the imperfect has *-u* the subjunctive has *-a* and the jussive has no overt ending. Wherever the imperfect has a disyllabic suffix (*-uuna, -iina, -aani*), the subjunctive and jussive moods have only the first syllable of this suffix. Otherwise, the suffixes do not alternate (*-na*).

For the time being, I will be discussing only the imperfect conjugation. The suffixal alternations will be treated later in the chapter.

Two aspects of the above paradigms will figure prominently in the following discussion and should be noticed right away. The first is that the suffix *-u* has a very irregular distribution: in particular it occurs in the singular except for the 1 pl form and the 2 f sg. Its occurrence in the 1 pl form is of particular interest. I will be arguing that *-u* is the default suffix and appears in the '1 pl' in virtue of the feature [pl] being realized at the prefix position by a special '1 pl' prefix *n-*, instead of at the suffix position as in the other plurals.

The second thing to notice is that all the feminine forms have the prefix *t-* except one, namely 3 f pl. I will be linking this to the fact that there is a special 'f pl' suffix *-na* occurring at the suffix position. I will show that because 'f' is realized at the suffix position, the regular feminine prefix *t-* does not occur in this form.

These two facts both represent what I will be calling *discontinuous bleeding*, whereby the realization of a feature at one position prevents (bleeds) its realization at another position. The existence of such phenomena argues for (1) a theory in which properties may be realized at various positions and not be linked invariably to a single position of exponence and (2) a theory in which properties have only one principal exponent, such that discontinuous bleeding is expressible as the discharge of the property only once as this exponent. To arrive at this theory, we will proceed very carefully through several alternative analyses, examining how each attempts to explain discontinuous bleeding.

1.2 Affixes and their Content

As a first step in morphological analysis, in (4) each affix of the imperfect is paired with the categories it is correlated with.

(4)	y-	3 m, 3 f pl
	-u	3 sg, 2 m sg, 1 sg, 1 pl
	t-	3 f sg, f du, 2
	-iina	2 f sg
	?-	1 sg
	-aani	du
	-uuna	m pl
	-na	f pl
	n-	1 pl

In a simple Word-Paradigm model, the statements in (4) exhaust the analysis of the conjugation. In other words, the correct morphological form can be generated by inspecting the category label and adding the affixes whose contents are a subpart of that label. For example, the 2 fem sing form *t-aktub-iina* has the affixes *t-* '2' and *-iina* '2 f sg', and no other affixes are compatible with the label. The analysis in (4) is therefore descriptively adequate, but deficient for a number of reasons.

The analysis in (4) misses several obvious facts. The first is that no verb has more than one prefix or more than one suffix. The second is that all the prefixes take the shape of a single consonant. The third is that the affix labels are highly complex, for example, *-u* represents four different sets of morphosyntactic properties, i.e. four homophonous affixes. The fourth is that certain feature combinations suspiciously recur, while others never appear. Consider the possible one-, two-, and three-feature combinations for affixes (certain feature combinations are omitted in virtue of impoverishment):

(5) 1, 2, 3, m, f, sg, du, pl, 1sg, 2sg, 3sg, 2du, 3du, 1pl, 2pl, 3pl, 2f, 2m, 3f, 3m, m sg, [3]m du, m pl, f sg, [3]f du, f pl, 2sg m, 2 sg f, 2pl m, 2pl f, 3sg m, 3 sg f, 3 pl m, 3 pl f.

Of the 34 possible affix labels, only 13 underlined ones are used, 2 (1pl and 1 sg) being used twice, giving a total of 15 different affixes. The features {sg f} group together in two labels (2sg f and 3sg f), but the features {2 pl} appear in none at all. A simple Word-Paradigm theory does not explain in any way why these 13 are used and not some other 13. We cannot appeal to parsimony: there are in fact only 11 different forms. A simpler system would have chosen 11 different affixes instead of 15, with a "maximal" differentiation of label types.

These considerations together suggest that the simple Word-Paradigm model is explanatorily inadequate.

1.3 The Extended Word-Paradigm Model

A considerable advance on the simple Word-Paradigm approach has been developed in the work of Anderson (1981, and following work), known as the Extended Word and Paradigm model (EWP). The two key observations of this approach are that affixes/rules sit in disjunctive blocks with the final rule possibly an Elsewhere rule (in the sense of Kiparsky 1973), which collects the 'unnatural' class of remaining categories. For example, the rules for suffixes in (4) can be expressed as in (6)

(6)
a. V, 2 f sg X --> X+iina
b. V, f pl X--> X +na
c. V, m pl X--> X+uuna
d. V, du X--> X+aani
e. Elsewhere, X --> X+u (in the imperfect)

The last rule, the elsewhere case, brings together the four "homophonous" *-u* affixes which as shown above do not form a natural class of any sort. The recognition that such Elsewhere rules exist is, in my opinion, the most important contribution of this line of research, since it reveals that there is a *system* underlying the group of affixes which may realize a given type of morphosyntactic property. In other words, knowledge of the inflectional morphology of a language consists in more than knowing a lexicon of affixes and the properties which they realize. Knowledge of morphology also must include principles which dictate the relative prominence among all the possible affixes (or rules) which may attach (apply) to a given stem.

In the EWP model, the fact that each verb has only one prefix and one suffix is expressed by having two blocks of rules, a suffix block and a prefix block. Rule blocks may correspond to position classes or templatic slots in other approaches, but need not be strictly isomorphic to these. Two expressions of the prefix block are given in (7) and (8), depending upon whether *y-* is treated as the elsewhere case (7f), or *t-* (8e):

(7)	a. V, 3 f pl	X--> y+X
	b. V, 1 pl	X--> n+X
	c. V, 1	X--> ʔ+X
	d. V, 2	X--> t+X
	e. V, f	X--> t+X
	f. Elsewhere	X --> y+X

(8)	a. V, 3 f pl	X --> y+X
	b. V, 3 m	X --> y+X
	c. V, 1 pl	X--> n+X
	d. V, 1	X--> ʔ+X
	e. Elsewhere	X --> t+X

It is important to observe that in either (7) or (8), the *n*-prefix rule (7b), (8c) precedes the ʔ- prefix rule (7c), (8d) since its environment is more specific (on the assumption that there is no [-plural] feature). This follows by the Elsewhere Condition and the similar principle identified with the Sanskrit grammarian Panini. Hence, the form **ʔ-aktub-uuna* cannot be derived from a 1 pl input representation, since 1 pl will condition *n*-prefix instead.

In terms of the ideas presented in the Introduction, we may say that the Elsewhere Principle discharges [1 pl] from the input representation. The domain of this discharge is this rule block only in this model, since disjunctivity requires that one rule in each block be chosen, but 1 and pl may figure again in conditioning rules of later rule rule blocks.

While this approach is significantly better than the simple Word-Paradigm model, certain embarrassments remain.

First, the fact that each prefix consists of one consonant only is a mere accident in this model. In fact there is a prosodic motivation for this fact: the imperfect stem consists of a skeleton of the shape VCCVC (or two heavy syllables), and this skeleton requires an onset, since, with few exceptions, full words in Arabic are consonant-initial. Yet this same requirement could be fulfilled by a prefix of shape CVC- or CVCVC- for example. Nothing in the EWP model prevents these unobserved occurrences. In fact, there is no relation between prosodic well-formedness and morphological rules in this model at all. The fact that the prefixing rules happen to satisfy the requirement that Arabic words have onsets is merely accidental.

Second, the suffix rules are unnecessarily complicated. Consider the 1st person plural for *n-aktub-u*. In order to prevent this form from appearing with the plural suffix *-uuna*, the *-uuna* rule (6c) is restricted to specifically *masculine* plurals. Since gender is not contrastive in the 1st person, this ensures that *-uuna* will never appear in the 1st person.

This solution, although it works technically for Classical Arabic, cannot work for the modern dialects which show similar behavior but have neutralized gender in all the plural forms, as I will show in section 1.13.1. Therefore, EWP requires some other means for preventing the affixation of *-uuna* in the 1 pl.

One possibility is to include among the suffix rules a specific rule which suffixes *-u* to the 1 pl form only. This rule, being more specific than the *-uuna* rule, will be disjunctive with it, and *-uuna* will not be suffixed to the 1pl form. The following illustrates this revised suffix rule block:

(9)

a.	V, 2 f sg	X--> X+iina
b.	V, 1 pl	X ---> X + u
c.	V, f pl	X--> X +na
d.	V, pl	X--> X+uuna
e.	V, du	X--> X+aani
f.	Elsewhere	X--> X+u (in the imperfect)

Again, this solution, while technically flawless, misses the point. Not only are there now two *-u* rules (9b) and (9f), but the above analysis fails to express the fact that the 1 pl form lacks the plural suffix *-uuna* precisely because the property 'pl' appears at the PREFIX position, together with the property '1' in the affix *n-* '1 pl.' Thus, the appearance of 'pl' at the prefix position BLEEDS it from the suffix position. Note that in the EWP framework this generalization can be expressed only if both the prefix *n-* '1 pl' and the suffix *-uuna* 'pl' belong to the same rule block. Nothing in principle prevents this, since rule blocks need not be confined to single positions of exponence. Yet this cannot be accomplished in this case. Both the prefix block and the suffix block have elsewhere rules: either *y-* or *t-* for the prefixes, and *-u* for the suffixes. If prefixes and suffixes belong to the same block, then there can be only one elsewhere rule. One cannot even say that, for example, the elsewhere rule of this combined block is a circumfix rule, adding, say, *y- -u*, since the domain of these two affixes is not coextensive: *y-* occurs without *-u* and vice versa.

A similar embarrassment arises in either treatment of the prefix block (7) or (8): both require two statements for the *y-* prefix rule. The source of the problem is the 3 f pl form, which fails to take the feminine prefix *t-*. To prevent this category from taking the *t-* prefix, it must undergo a *y-* prefix rule preceding the *t-* prefix rule. Because this *y-* prefix rule must precede the *t*-prefix rule, the *y-* prefix rule cannot be the Elsewhere rule. Hence the fact that the categories which take *y-* prefix (3 f pl & 3 m) do not form a natural class forces there to be two *y-* prefix rules as in both (7) and (8).

Again, the fact that there are two *y-* prefix rules misses an obvious point, namely that the reason the 3 f pl form does not take the prefix *t-* 'f' is because f is realized at the suffix position along with pl as *-na* 'f pl'. Suppose again then that the disjunctivity of *t-* and *-na* can be expressed by putting both in the same rule block. As in the case of *-uuna* vs. *na-*, this leads to difficult problems.

The following partial rule block illustrates the putative rule block in which *t-* 'f' and *-na* 'f pl' are disjunctive:

(10)
a. f pl ---> X+na
b. 1 pl ---> n+X
c. 1 ---> ʔ+X
d. f ---> t+X

Now, the rule

(11) 2 ---> t+X

cannot go in the block in (10), since it is not disjunctive with *-na:* as shown by *t-aktub-na* 2 f pl. Therefore the *y*-prefix rule cannot be the Elsewhere rule in block (10), for if it were, then all 2nd person arguments would have prefix *y-*, contrary to fact. We are thus still left with two separate *y-* prefix rules, since the *y*-prefix rule still cannot be expressed as the elsewhere case.

In the EWP model as exhibited so far, we cannot even express the disjunctivity of *t-* 'f' and *-na* 'f pl' at all, as I will now show. Since *-na* is disjunctive with *-u*, which is an elsewhere rule, then *-u* must be the elsewhere case in block (10). But *-u* is disjunctive with all the suffixes, including dual *-aani*. So dual *-aani* must also go in block (10). But

-aani is not disjunctive with *t-* 'f', as evidenced by the 3 f dual form *t-aktub-aani*. But we have already ascertained that *t-* 'f' must go in block (10) since it is disjunctive with *-na*, but clearly it cannot go in block (10) since it is not disjunctive with *-aani*. This is the paradox. The rule-based approach must retain two *y*-prefix rules and treat the disjunctivity of *t-* 'f' and *-na* 'f pl' as an accident. In other words, the two *y-* affixes are held to be accidentally homophonous.[1]

EWP must treat all such cases of bleeding across rule blocks through duplication of rules (or identity rules, as I will show in section 1.13.1). This provides the clue that EWP is missing an important generalization.

Anderson (1986), in a discussion of nearly identical facts from Georgian, modifies the conception of disjunctivity which I have so far attributed to the EWP model. While this modification solves the problem in Georgian, it is insufficient for Arabic. The modification can be stated in two parts:

(12) two rules A, B will be disjunctive *across rule blocks* if
(1) B refers to a subset of *features* referred to by A and
(2) A precedes B.

This can be schematized below:

(13) Rule A: +F +G X --> X′
Rule B: +F X --> X″
A applies, B is blocked, regardless of rule block location

Note that Rule A must precede Rule B in order for there to be a disjunction; if Rule B applies first, then rule A may also apply. (If this were not true, multiple exponence would be entirely impossible, as I will show in the next section.)

For example, the *n-* rule refers to '1 pl' and the *-u* rule refers to 'pl.' Therefore these two rules are disjunctive and only *n-* '1 pl' applies, even though the two rules must sit in separate blocks, as has been established. This seems to solve the problem of discontinuous bleeding.

---

[1]If 3 f pl did take the *t*-prefix characteristic of feminines, then the 3 f pl form would surface *t-aktub-na* and be homophonous with the 2 f pl form. A functional explanation might be advanced as to why these two do not neutralize: it would be too ambiguous. Aside from the vagueness of such notions as "too ambiguous" this argument falls apart because the 3 f sg and 2 m sg forms are in fact homophonous. Why is this homophony permitted, while the other, putative homophony between 3 f pl and 2 f pl is not? One cannot say. The functional explanation is less interesting insofar as it cannot answer this question. Furthermore, as I show later, precisely this neutralization did occur in two separate branches of Semitic: Hebrew and S. Arabian, so such an ambiguity cannot be ruled out in principle.

However, only one of the two cases of discontinuous bleeding can be handled by this revision in the definition of disjunctive rule application. The bleeding of *t-* 'f' by *-na* 'f pl' would seem to follow straightforwardly as well: the two rules should be disjunctive, and a 'f pl' input should undergo only the *-na* rule. However, the more specific rule must precede the less specific rule by the second criterion of (12). Since the *-na* rule goes in the suffix block, this implies that the suffix block must precede the prefix block. This further implies that the *-u* 'pl' rule must precede the *n-* '1 pl' rule. Yet if this is so, then these will not be in a disjunctive relation, contrary to fact. The resulting paradox of ordering establishes that, insofar as both of these disjunctions are real, they cannot both be captured by Anderson's (1986) proposal.

Before leaving this topic, I should like to point out that (12) does not follow from the original conception of the Elsewhere Principle, but is in fact somewhat stipulative. The Elsewhere Principle requires merely that the SD of one rule be contained in the SD of another. Consider the following rule system:

(14)	Block I:	+F +G	X --> A
		Elsewhere	X --> B
	Block II:	+F	X --> C
		Elsewhere	X --> D

The condition on rule disjunction across blocks in (12) must explicitly require that rule B will be blocked by rule A only if rule B refers to some *nonempty* set of features which is a subset of the features referred to by rule A. Thus the X-->A rule will block the X--> C rule, but *not* the Elsewhere rule of Block II, since the Elsewhere rule of Block II refers to no features. On the original conception of the Elsewhere Principle, if disjunction were to apply across rule blocks regularly, then the X-->A rule would block the Elsewhere rule of Block II, just as it blocks the Elsewhere Rule of Block I. Yet this clearly is undesirable: hence (12) must require the blocked rule to refer to a subset of features of the blocking rule. This added stipulation reveals that disjunctions within rule blocks (within the domain of a single Elsewhere rule) and disjunctions across these domains should not be captured by means of a unitary principle such as the definition of disjunctive application.

Developing a theory in which such discontinuous bleeding effects can be captured directly and efficiently is the topic of this chapter and one of the main objectives of this thesis. I will show that these discontinuous bleeding effects can be captured directly by capitalizing on the difference between primary and secondary exponents, as discussed in the Introduction, and in a theory in which rules discharge both positions and features independently.

1.4 Allomorphy vs. Discontinuous Bleeding

In order to capture discontinuous bleeding, let us suppose that a feature can figure only once in any morphological rule, i.e. it may be discharged only once. This move requires in effect that each morphosyntactic property have one and only one exponent.

Such a move is too radical, as it ends up permitting no allomorphy whatsoever. Consider for example the form below:

(15) t-aktub-iina 'you (f, sg) write'
2-write-2 f sg

The prefix *t-* expresses '2,' but the suffix is also limited to the 2nd person. For example, the 3rd person feminine form is *t-aktub-u* with the default imperfect suffix *-u*. Therefore *-iina* cannot express merely 'f' but must express '2 f' at least.

By hypothesis, a feature may be discharged only once. Therefore, if *-iina* must be '2 f' then '2' cannot condition the prefix *t-* in *t-aktub-iina*. The expected form is then **y-aktub-iina*. Yet this is incorrect.

Because of such examples, which are extremely numerous, proponents of paradigm theories insist that the discharge of a feature is limited only to the rule block domain: '2' must appear both in the *t-* rule block and the *-iina* rule block. Therefore, it is argued, features are never used up nor discharged. All disjunctivity must be a consequence of rule disjunctions, including in some cases and in some versions of the theory, disjunctions applying across rule blocks as discussed in the last section.

I argued in the previous section that the idea that features are not discharged except via rule disjunction makes it impossible to express discontinuous bleeding effects, which in general must be treated as anomalies in any paradigm theory. I showed that the domain of the elsewhere rule for each block was not the same as the domain of the necessary disjunctions, and that permitting certain disjunctions across rule blocks was still insufficient. For the sake of argument, I was then led to assume that the discharge of a feature was not limited to the rule-block domain, but instead was absolute and irrevocable. Yet the facts of allomorphy as in *t-aktub-iina* seem to demand that always invoking discontinuous bleeding is too powerful.

The solution to this apparent paradox lies in the idea of principal vs. secondary exponents. The claim advanced here is that if an affix is the principal exponent of a feature, then the rule introducing/spelling-out the affix discharges the feature. Discharged features continue to be visible, but may thereafter condition only allomorphy in rules which discharge other features.

To show this concretely, consider again the form 2 f sg *t-aktub-iina.* This form will be subject to two rules, given below:

(16) a. 2 t-
b. f (2) -iina

The first rule discharges '2' and fills the prefix slot selected by the imperfect stem with the segment *t-*. The second rule discharges 'f,' but applies only to forms in which '2' has previously been discharged. This allows allomorphy to be expressed. In other words, the suffix *-iina* is analyzed as the allomorph of 'feminine' that appears in the 2nd person.

I will use the parentheses notation to indicate those properties which must be previously discharged in order for a given rule to apply. When a feature has been discharged, it follows that some affix realizing that feature appears in the string. Therefore a rule which applies only when a certain feature has been discharged will realize an affix only on a string in which this feature has already appeared. Thus, we may speak of the discharge of a feature as the process which makes that feature visible in the string. For example, (16b) will apply only to a string in which the feature '2' has been made visible in virtue of its appearing on some affix in the string as below:

(17) $[t_2]$ - Stem

After the affixation of *t-* '2,' the feature '2' is discharged and therefore is visible in the string to rule (16b).

On this view, features (like inflectional class or gender or animacy on nouns) which are properties of the stem will always be visible in the string and therefore may always condition allomorphy rules for any affixes.

Turning now to cases in which discontinuous bleeding is desirable, consider the 1 pl form *n-aktub-u.* Of the rules below, only (18a) will apply, since (18b) and (18c) cannot apply once '1 pl' is discharged.

(18) a. 1 pl n-
b. 1 ?-
c. pl -uuna

The rules in (18) are automatically orderr ;ording to complexity, and the hierarchy of features (person > number).

In terms of principal and secondary exponents, an affix (rule) will have be a principal exponent of those features which are discharged by the affix (rule); the affix will be a secondary exponent of any other features, which by convention I will write in parentheses. For example *-iina* 'f (2)' is the principal exponent of 'f' and a secondary exponent of '2' in the form *t-aktub-iina*. The conditioning environment can be read "f *in the environment of* 2." In other words, *-iina* is the allomorph of 'f' that appears in a form in which '2' has been discharged.

It is often assumed that dividing exponents into principal and secondary has no basis in linguistic reality. Especially in the domain of inflectional morphology, any given analysis is untestable-- so the complaint goes -- since inflectional word formation is relatively restricted in the number of different generable types (i.e. the categories of a paradigm are a closed class). This objection can be made only because the principal/secondary division is not obvious and is dependent upon an analysis of a morphological system as a WHOLE -- just as the analysis of the phonology of a language consists in more than a taxonomy of phones and phonemes. The conclusion that *-iina* is a secondary exponent of the property '2,' can be arrived at only after it is deduced that *t-* is the primary exponent of '2,' a conclusion that is dependent upon the analysis of all the other affixes of the system.

It should be borne in mind, in fact, that any deviation from the naive paradigm model I presented in 1.2 -- where each affix is simply listed with those categories in which it occurs -- entails certain choices. For example, any analysis which postulates an elsewhere rule or affix is based on some evaluation of which classes of properties constitute natural classes and which do not, any analysis which isolates stems which figure in subsequent rules has already performed at least one morphological segmentation, etc. Resistance to further analysis on the grounds that it has no provable linguistic reality amounts to nothing more than a pre-generative prejudice.

A useful criterion by which a morphological analysis may be tested, however, is historical change. Historical changes should be easily expressible as changes in parameter settings, loss or gain of rules, and simple reanalysis of representations. I will be focusing on the reflexes of the prefix-paradigm throughout the Afroasiatic family in sections 1.12 through 1.18, showing how historical changes in fact support the present analysis.

Finally, it should be noted that the processes of 'discharge' and 'secondary exponence,' as shown here, are entirely analogous to 'filling categorial signature' and 'subcategorization of previously percolated features' in a strict lexical model such as Lieber (1989, 1992), except IN REVERSE. The difference is that the model advanced here generates (inflectional) morphology from morphosyntactic representations of freely composed feature matrices, discharging the features contained in these representations. Strict Lexicalism, in quite the reverse process, builds up morphosyntactic representations by percolating features from freely added affixes. Therefore, it is an obvious question whether Strict Lexicalism is not in fact correct. In the next section, I review this approach and reasons why I reject it.

1.5 Affix-based Models

In this section I discuss one possible analysis which falls under the heading morph-based or affix-based, as opposed to the word-paradigm models presented in the previous section. We begin with an analysis in the spirit of Lieber (1980, 1989, 1992), showing that these theories also fail to capture directly and unproblematically the interesting generalizations of the Arabic prefix conjugation.

Lieber's (1980) theory treats affixes as lexical entries which combine together freely, subject to each affix's subcategorization restrictions. In this way, the syntax of affixes and roots is formally parallel to the syntax of words as proposed in Chomsky (1981). Lieber's (1992) theory abandons lexicalism entirely and situates all word-formation within syntax. The syntax of affixes and roots in this more radical model IS the syntax of words. Both models retain an essential idea, namely that word-pieces are listed elements which combine freely subject to principles of selection or subcategorization and a general algorithm of concatenation such as move-α.

In (19) I again list each affix in the Arabic prefix conjugation conjugation:

(19)

?-	1
t-	2
t-	f
-iina	f, [2 + ___]
-aani	du pl
n-	1 pl
-na	f pl
-uuna	pl
y-	3 (?)
-u	imperfect (?)

The left column shows the affix, and the right column shows the minimal feature representation each affix must have. Observe first that the occurrence of the suffix *-iina* only in the 2nd person is captured in this framework by giving the suffix a subcategorization frame requiring it to attach only to items bearing the feature '2.' The only way for an item to gain the feature '2' is by prefixation of *t-* '2.' Therefore such nonoccurring forms as **y-aktub-iina* are never generated.

Several areas of uncertainty arise immediately in the affix-based analysis.

The first problem arises in the analysis of elsewhere or "empty" affixes. For example, the feature-content of the *y*-prefix is unclear. Recall that *y-* appears in 3 m and 3 f pl, but does not appear in the 3 f sg or 3 f dual. Compare the following forms:

(20)

a.	y-aktub-u	3 m sg	'he writes'
b.	t-aktub-u	3 f sg	'she writes'
c.	t-aktub-aani	3 f dual	'they (f, du) write'
d.	y-aktub-uuna	3 m pl	'they (m, pl) write'
e.	y-aktub-na	3 f pl	'they (f, pl) write'

As I showed in the previous section, the *y-* prefix is the elsewhere case.

Similarly, the suffix *-u* is another elsewhere case. The categories where *-u* appears do not form a natural class: 3 sg, 2 m sg, 1 sg, 1pl.

In any affix-based theory, affixes combine subject to their subcategorization frames. The features carried by each affix percolate from the affix to the word as a whole (subject to certain restrictions such as headedness and the categorial signature). Therefore, any affix-based analysis must identify what exactly are the features which are borne by each affix as part of its lexical entry. In the case of elsewhere affixes like *y-* and *-u* this problem is acute.

For the time being, let us suppose that *y-* is merely '3' and *-u* is 'imperfect.'

Second, Lieber's theory (or any affix-based theory) must capture the position-class phenomena in some way as well. If affixes were to freely combine, then it would be possible with the affixes in (19), for example, to derive such forms as:

(21)	a.	*t-t-aktub-uuna-u 2 -f-write-pl -IMPF	2 f sg
	b.	*y-t-aktub-aani-u 3-f-write-dual-IMPF	3 f sg
	c.	*t-aktub-iina-uuna-u 2-write-f-pl-IMPF	2 f pl

The affix-based theory's solution is to make each affix's subcategorization frame specify its environment of occurrence. Hence, all the prefixes will, for example, select the stem VCCVC as immediate sisters, as in (22):

(22)	ʔ (1)	[___	$[_V$ V..
	t (2)	[___	$[_V$ V...
	t (f)	[___	$[_V$ V....
	n (1 pl)	[___	$[_V$V....
	y (3)	[___	$[_V$ V....

No two prefixes may combine, since once one has attached, the subcategorized environment (strictly stem-adjacent) has been used up. If two affixes belong to the same position class, then their subcategorization frames specify the same location.

Lieber (1989) makes extensive use of this sort of subcategorization in a treatment of the agreement system of Yavapai. Each position in which an affix may occur is represented as either another level of derivation in the sense of the Lexical Phonology Model of Kiparsky (1982) or Mohanan (1986), or as another bar-level of the verb: V, V′, V″.

A similar story can be devised for the Arabic suffixes. For each of these, however, the subcategorization frame must specify $[\,[_V{}^1]$ ___ $]$, where the superscript '1' (or some equivalent to it) means a verb stem which has already been prefixed on the previous stratum.

Position classes may therefore be subsumed into lexical strata or their 'syntax-only' equivalent, namely X-bar levels.

There are two objections to this. The first is that the evidence motivating the existence of level-ordering has been challenged since lexical phonology was first proposed. For example, Fabb (1988) carefully reviews the affixes of English and concludes that level-ordering in English is is an illusion.

Second, if level-ordering is abandoned and only selectional restrictions and 'potentiation' of certain affixes by others are retained, as Fabb suggests, then the fact that each subcategorization frame must repeat the *same* environment for each prefix and each suffix shows that this approach is missing a generalization. It is an accident of the subcategorization frames that prefixes are level I and suffixes are level II in the Arabic prefix-conjugation. In other words, there is no principled reason why there could not be a level II prefix in this conjugation. The fact that all the verbs of this conjugation have one prefix and one prefix only is treated as a sort of epiphenomenon arising from a conspiracy of many subcategorization frames, or equivalently, the level at which each prefix may attach. This problem does not disappear in Lieber's (1992) theory in which the relevant levels are not strata in the sense of lexical phonology but rather bar-levels in an expanded system of syntactic projection (pp. 35ff.) Rather, the problem of stating the same privileges of occurrence again and again is merely shifted to the syntactic domain.

Lieber's levels essentially enforce a disjunctivity among a certain group of affixes in the same way Anderson's disjunctive blocks do. Like rule blocks, subcategorization does not limit disjunction to any particular location of exponence. But whereas the paradigm model cannot express discontinuous bleeding, the affix model simulates discontinuous bleeding by means of filling 'categorial signatures.'

To understand how this works, consider now the eight bad forms in (23), (24) and (25). Note that each of these words can be formed using one prefix and one suffix, given the affixes as I have defined them.

(23)	a.	*t-aktub-na	f-write-f pl
	b.	*n-aktub-uuna	1 pl-write-pl
(24)	a.	*?-aktub-aani	1-write-dual pl
	b.	*n-aktub-aani	1 pl-write-dual pl
	c.	*?-aktub-na	1-write-f pl
	d.	*n-aktub-na	1 pl-write-f pl
(25)	a.	*?-aktub-uuna	1-write-pl
	b.	*t-aktub-uuna	f-write-pl

(Note: The ill-formed cases (23a) and (25b) are homophonous with well-formed words with *t-* '2' prefix.)

The two bad forms in (23) are precisely those cases where I argued in the last section that discontinuous bleeding applies. In an affix-based theory, some means must exist to prevent such forms.

Lieber's solution in this case, as was outlined in the Introduction, is to propose that each word has a 'categorial signature,' a matrix of feature value positions which may be filled by features percolating from affixes. For our purposes, the categorial signature of an Arabic verb might appear something like this:

(26)
$$\begin{bmatrix} \_\_\_\ 1 \\ \_\_\_\ 2 \\ \_\_\_\ 3 \\ \_\_\_\ \text{dual} \\ \_\_\_\ \text{pl} \\ \_\_\_\ \text{f} \\ \_\_\_\ \text{imperfect} \end{bmatrix}$$
etc.

Lieber assumes that morphosyntactic features are binary, but this does not affect the general point.

Now it is crucially the case that those features which are represented in the categorial signature may have their values supplied only once. Returning to the two forms in (23), observe that in both cases the same feature appears in the suffix position as occurs in the prefix position. In an affix-based model, assuming that prefixes attach first, then the prefix features will percolate to the signature and fill the slot. Thereafter, no affix may attach which would percolate that feature. For example, in the form *n-aktub-uuna*, the prefix *n-* '1 pl' will percolate the features '1 pl' and fill these slots in the signature. The suffix *-uuna* is barred from attaching since it supplies the feature 'pl' again.

In this way, the device of filling the categorial signature allows the effects of discontinuous bleeding to be expressed, preventing the bad forms in (23). From this point of view, the affix model fares better than the paradigm model.

Yet this solution betrays a fundamental peculiarity of the affix-model. Strict lexicalist or affix-only approaches in principle permit no distinction between inflection and derivation: both are subject to the same principles of word-formation. Nevertheless, the fact that some features compose a 'signature' while others do not shows that the signature-feature supplying affixes and derivational affixes are different after all. For it is clearly not the case that all features fill a signature.

Canonically derivational affixes may supply category-changing features, indeed, may supply an entirely new categorial signature with a new category. For example, the suffix *-able* carries the feature [+N +V] (i.e. adjective) and attaches to [-N +V] (i.e. verb) categories to form deverbal adjectives such as *eat-able*. The features borne by the affix are already represented by the stem to which it attaches. Not only does the affixation of derivational material not obey the 'categorial signature' behavior ([+V] may attach to [+V]), it also permits feature-changing as well ([+N] attaches to [-N], and changes the value of the whole word to [+N]). This entails that those affixes such as English *-ing* which are allegedly both inflectional and derivational must in fact be sets of homophonous affixes, some with and some without categorial signatures and feature-changing behavior.

The second group of bad forms, (24) requires a different explanation (or are amenable to one, in the case of (24b)). These forms can be blocked by the same sorts of filters I proposed in the discussion of Impoverishment in the Introduction:

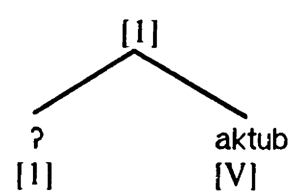
(27)	a. *?-aktub-aani	1, dual pl	*[1 dual]
	b. *n-aktub-aani	1 pl, dual pl	*[1 dual]
	c. *?-aktub-na	1, f pl	*[1 f]
	d. *n-aktub-na	1 pl, f pl	*[1 f]

Recall that in an affix-based theory, there is no input representation to Morphology which conditions the realization of affixes, and which may be appropriately Impoverished to derive the correct neutralizations. The morphosyntactic representation of a word is derived from the features contained in its constituent parts. Some of the features of the parts are inherited by the whole by means of Percolation.

The filters then must act on the product of word-formation rather than on the input to word-formation. Such filters will be called Output Filters.

For example, suppose the prefixes attach first and ?- with feature [1] attaches freely to the stem, giving (28):

(28)



The feature [1] percolates from the affix ?- to the whole. Consider the ill-formedness of (27a,b). Here, the obvious generalization is that the dual suffix *-aani* cannot attach to forms which have a 1st person prefix, either ?- '1' or *n-* '1 pl.' When *-aani* 'du pl' suffixes to these forms, then [dual pl] will percolate to the categorial signature. The resultant matrix, [1 dual pl], is barred by the output filter *[1 dual]. Similar facts hold of the [1 f] forms (27c,d).

It is however technically possible to capture the generalizations expressed by output filters through more complicated subcategorization frames. For example, *-aani* and *-na* would require the frames in (29):

(29)	*-aani* 'du pl'	[[V, -1] ____]
	-na 'f pl'	[[V, -1] ____]

This move is unsatisfying because the feature [-1] needs to be invoked only to express this restriction, which is in fact a general property of Arabic: feminine is never distinguished in the 1st person, nor is dual. This generalization is not at all the property of individual affixes. Stating this general fact over and over in a number of subcategorization frames again loses the obvious generality of the facts. I will show in section 1.14 that treating *-aani* as a specifically [-1] suffix also complicates the historical extension of this suffix in the S. Arabian language Meħri.

Output filters are necessary in affix-based theories for another reason. The categorial signature of a given word will not be completely filled by the features percolated from its affixes. To give one example, the form *t-aktub-u* is 2 masc sg, but the content of its affixes is merely '2' for the prefix and 'imperfect' for the suffix. (Observe that *t-* cannot be '2 masc' since *t-* also appears in the 2 f sg form *t-aktub-iina*, for example.) We may assume, following Lieber's discussion (1992:101), that the remaining features, 'masc' and 'sg,' are supplied by default rules operating at the end of the derivation. If 'masc' is filled in automatically, then some provision must be made for preventing it from being filled in for the 1st person form *ʔ-aktub-u*. Output filters must be invoked to block the redundancy rule filling in gender.

Since output filters must be incorporated into the affix-based theory, forms like (24) are filtered out after they are put together. Output filters makes clumsy subcategorization frames like (29) unnecessary and serve the same purpose in the affix-model as do Impoverishment filters in a paradigm-model.

Outut filters cannot explain the ill-formedness of the two remaining bad forms in (25), **t-aktub-uuna* 'f-pl' and *ʔ-aktub-uuna, '1 pl.' An output filter *[f pl] cannot be invoked to prevent the derivation of **t-aktub-uuna* 'f pl', since [f pl] must be allowed to combine, as for example in the suffix *-na* 'f pl.' Similarly, no output filter of the form *[1 pl] can be invoked to block *ʔ-aktub-uuna, '1 pl,' since the prefix *n-* is itself '1 pl.'

This is where the affix-based model really runs into trouble. Such forms must be prevented by recourse to Blocking, as first discussed by Aronoff (1976). However, even appealing to Blocking will not work here, owing to an ordering paradox to be discussed.

First, however, let us be explicit about what is meant by Blocking. The following formulation is given by Jensen (1990:90), following Kiparsky (1983:15).

(30) Avoid Synonymy Principle
The output of a lexical rule may not be synonymous with an existing lexical item.

The typical example of such a restriction is the blocking of **glori-os-ity* by pre-existent *glory*. The word *glory* is necessarily pre-existent since it is listed lexically, whereas *glori-os-ity* is derived by at least two derivational processes. However, what is significant about blocking as a phenomenon is that it serves only to make blocked forms *dispreferred* but not strictly ungrammatical, and when used, they are forced to acquire residual meanings. Since *glory* is an existing lexical item, blocking ensures that *gloriosity* can be used only in a sense not already associated with *glory*. Blocking does not entail that *gloriosity* is ungrammatical. Aronoff (1988) suggests, following Horn (1984), that Blocking is 'probably explainable in terms of general pragmatic principles ... which extend to purely syntactic cases like Avoid Pronoun... (p. 767).'

Consider the differences between *cook* and *cook-er*, one used of a person, the other only of machines. If Blocking is truly the principle which is at play in ruling out such forms as *ʔ-aktub-uuna, '1 pl,' then we should expect that ʔ-aktub-uuna would merely acquire some subsidiary meaning, some special kind of 1st person plural not covered by n-aktub-u '1 pl.' Yet this never happens.

Leaving aside this objection for the moment, let us consider how Blocking might be made to prevent the unwanted forms. It now seems that all that need be assumed is that ** ʔ-aktub-uuna* is blocked by pre-existing *n-aktub-u*, and that **t-aktub-uuna* is blocked by pre-existing *y-aktub-na.*

Recall that prefixation occurs at level I and suffixation at level II. The blocking of ** ʔ-aktub-uuna* follows straightforwardly. *ʔ-aktub-uuna* is blocked at Level II by *n-aktub-u* which has already been derived at Level I:

(31)	Level I:	n-aktub	'1 pl'	ʔ-aktub	'1'
	Level II:	n-aktub-u	'1 pl impf'	-uuna suffixation blocked	

On the other hand, the blocking of **t-aktub-uuna* cannot be derived, since the form that would block it is derived at the same level, level II. The suffixation of *-na* and the suffixation of *-uuna* occur at the same level, so nothing in principle requires *y-aktub-na* to be 'pre-existing' at the time when *-uuna* is 'trying' to attach to *t-aktub-*.

In a model such as Lieber (1992) where word-formation takes place in syntax rather than the lexicon, it is unclear how blocking could possibly be at work. Aronoff (1988) points out that while words may block words and words may even block phrases (so blocking is not limited to the word-domain), the blocking item is always a listed item and the blocked item is a derived form. But the blocking of *ʔ-aktub-uuna by pre-existing n-aktub- is not the blocking of a derived form by a listed form, but rather of a derived from by a previously derived form.

In the paradigm model, such a problem cannot arise, since the input representation [3 f pl] will condition *-na* 'f pl' and never *-uuna* 'pl' by the Paninian principle. This principle has no analogue in an affix-based model. Although blocking attempts to simulate the Paninian principle, in fact it may do so only when an ordering paradox does not arise.

In general, it is the absence of such rules of 'priority' which will continually hamper the affix-based model.

In order to evade this paradox, the affix-based model could treat *-uuna* as specifically masculine plural. This move, while getting the forms right for Classical Arabic, will not work for those modern Arabic dialects which have lost gender in the plural, as will be detailed in section 1.13, and therefore must be abandoned in principle.

Alternatively, the prefix *t-* could be specifically nonplural (unless dual). While this would again capture the facts, it is certainly stipulative: the distribution of *t-* 'f' would have to be limited to precisely those environments where a more specific feminine affix (*-na* 'f pl') exists at a *later* level of affixation.

Now to compare the different models, consider how each answers the question of why the different forms in (23), (24) and (25) cannot occur. Some forms are ruled out for more than one reason. The full set of ill-formed verbs is reiterated in (32):

(32)

			Paradigm Theory	Affix Model
a.	*t-aktub-na	f, f pl	Special *y-* rule	Categorial Signature
b.	*n-aktub-uuna	1 pl, pl	Special *-u* rule	Categorial Signature
c.	*ʔ-aktub-aani	1, dual pl	Filter on input	Filter on output
d.	*n-aktub-aani	1 pl, f pl	Filter on input	Filter / Categorial Sig.
e.	*ʔ-aktub-na	1, f pl	Filter on input	Filter on output
f.	*n-aktub-na	1pl, f pl	Filter on input	Filter / Categorial Sig.
g.	*ʔ-aktub-uuna	1, pl	Elsewhere Condition	Blocking??
h.	*t-aktub-uuna	f, pl	*y-* rule and E.C.	Special feature content

Several important observations come from this table. The first observation is that both theories must have recourse to filters of some sort.

The second observation is that where the affix-model needs to appeal to Blocking, the paradigm-model model cannot generate the bad form owing to the Elsewhere Condition.

The final observation is that where the affix-model invokes the Categorial Signature, the paradigm model must have more complicated rules, i.e. the affixes themselves are taken to express more morphosyntactic properties so as to restrict their appearance to categories in which they will not lead to multiple exponence. Constraints on multiple exponence are accordingly accidents for the paradigm-model, principled for the affix-based model.

Given that the affix-based model can simulate the effects of multiple exponence in virtue of subcategorization and allomorphy, its claim must be only that multiple-exponence effects are less probable, i.e. more marked. Paradigm-theories, on the other hand, because they are unable to directly express discontinuous bleeding, treat multiple-exponence effects as unmarked phenomena. Discontinuous bleeding entails more complicated rules, such as the special rule -- V, 3 f pl --> y+V -- and so discontinuous bleeding is marked. But bleeding within a position (*n-* '1pl' bleeding ʔ- '1') is principled and inviolate. Affix-based models in turn have difficulty enforcing bleeding within positions, since essentially nothing "forces" one affix to be chosen instead of another: affix-based models cannot invoke the Paninian principle directly, but must appeal to Blocking. Yet sometimes Blocking cannot work, e.g., the word **t-aktub-uuna* (3 f pl) is ill-formed and, owing to an ordering paradox, cannot be blocked at all.

It is worth pointing out that both theories can generate all the correct forms, either with special rules or extra features or subcategorizations. This debate is not a question of *descriptive adequacy*, in the sense of Chomsky (1965:25ff.) Rather, both the affix-based and the paradigm-based analyses fail to directly express certain generalizations about the Arabic prefix conjugation, generalizations which any theory of morphology should explain directly, and in consequence each has recourse to remarkably similar maneuvers to deal with the facts.

In the next section, I turn to a third model, Halle's abstract-morpheme model (1989a, 1990, 1992). A thorough examination of the underpinnings of his treatment of positions as things in themselves (by means of the abstract Q-strings which occupy them) will lead to a new theory which captures the relevant generalizations more directly, while allowing the possibility of a more direct mediation between syntactic processes and morphological ones.

1.6 Abstract Morpheme Models and Word Templates

One of the most striking facts about the so-called prefix-conjugation in Arabic (which includes the imperfect, subjunctive, and jussive) is that every verb has at least and at most one suffix and at least and at most one prefix. This fact can be represented by the device of a template to which all imperfect verbs must conform:

(33) [[Affix Stem] Affix]

In Selkirk's (1982) model of the internal syntax of words, such a template might be constructed through a phrase structure rule as in (34):

(34) Verb (imperfect) ---> Affix Root Affix

The phrase marker so generated serves as the template and affixes and roots are then inserted into this tree. No imperfect verb could have more than one prefix or suffix, since affixes may be inserted only into trees generated by the phrase-structure rules of the lexicon. Things then proceed much as in Lieber's affix-based model, but without the subcategorization frames expressing strict Stem-adjacency.

On such a view, the internal syntax of words is entirely a product of the phrase-structure rules of the lexicon. However, much work in syntax since Selkirk's proposal has taken the position that the syntax of complex words is at least in part determined by syntactic affixation. These views are most extensively discussed in Fabb (1984) (for derivational morphology of English) and Baker (1988) for grammatical function-changing morphology in a number of languages. The hypothesis that principles of syntax (specifically, Case theory and the Head-Movement Constraint of Travis (1984)) constrain the operation of syntactic affixation has reintroduced the idea that the syntax of affixes is reducible in many cases to the syntax of X^0 projections. The English auxiliary Affix-Hopping and Do-Support Rules first discovered by Chomsky (1957) are typically viewed as syntactic affixations in much current work (Pollock 1989, Chomsky 1989): these still represent the most compelling evidence that Head-Movement is implicated in the construction of inflectional (as opposed to grammatical function-changing) morphology.

Because this view has been exceedingly influential in recent years, I will devote some energy to making absolutely clear what is assumed by the idea of syntactic affixation.

The first question to be asked concerns the product of the Head Movement rule which is implicated in syntactic affixations. When X^0 head-moves to adjoin to Y^0, the result is an adjunction structure as in (35):

(35) $[_{Y^0}\ Y^0\ X^0\ ]$

Because syntactic structure is unlinearized (Marantz 1988), X^0 is either left-adjoined or right-adjoined to Y^0, depending on the subsequent linearized interpretation of (35) at Phonological Form.

There is a striking formal similarity between an adjoined structure such as (35) and a root-affix phrase-structure in the sense of Selkirk (1982), illustrated in (36):

(36) $[_{\text{Affix}}$ Affix Root]

This similarity has led to the idea that the two are in fact identical: that the syntax of X^0 projections IS the syntax of roots and affixes, or that the two differ only trivially, perhaps by the addition of extra affixes in morphology that are not specifically represented as syntactic atoms. Obviously, this is the simplest hypothesis of the mapping relation between s-structures in syntax and morphological forms, and can be embodied in slogan-form: "Every affix/root is a syntactic head."

Precisely this conclusion is implicit in analyses where morphemes are implicitly conflated with X^0s. I exhibited one such analysis (Giorgi & Pianesi 1991, to appear) in section 0.2.1.

Halle (1989a, 1990) and following work has incorporated a part of this attitude towards inflectional morphology in a model which takes the output of syntax, i.e. structures of the sort (35), as the inputs to morphological rules which constitute a component of the mapping to Phonological Form. This component is called Morphology.

As I showed in the Introduction, Halle's template for nouns in Latvian, for example, appears in (37):

(37) [[[Root] Q, Theme] Q, Number-Case]

Each Q serves as a dummy occupying a string position until it is rewritten as a phonological string by a Spell-Out Rule. Besides its location in the string, each Q is also identified by a label. The first Q in (37) is identified as *Theme*. The second Q is identified by some collection of Number and Case features. Each Q and its corresponding identification serves as a bipartite entity, the whole of which is an *abstract morpheme*.[2]

It is important to recognize that the representation in (37) is that of a *type*. In other words, all nouns conform to this type, but (37) represents no noun in particular. A particular instantiation of the type in (37) would be:

(38) [[[*gov*] Q,Theme] Q, Plur-Acc] ---> *gov - i - s* 'cows-ACC'

The most natural question is: what guarantees that all Latvian nouns conform to the type expressed by the representation in (37)? Specifically, if Q,Theme is not a syntactic head, then we cannot appeal to Head-Movement in syntax to construct ALL of (38). The presence of an obligatory affix Q,Theme in all nouns, on the other hand, is such a general fact that it seems appropriate to capture it immediately by a template such as (37).

As I argued in section 0.2.1.1, the syntax of affixes and roots within a word is not ultimately derivable *entirely* from syntactic affixation. Words have autonomous, extra-syntactic structure. We must revise our slogan to "Some affixes are (isomorphic to) syntactic heads."

This conclusion still leaves us with the question of how to ensure that all nouns in Latvian comply with the template in (37).

Turning now to Q,Number-Case, we can identify two facts of interest. The first is that Number properties and Case properties always show their exponence together. In other words, as in most Indo-European languages, the number and case are "fused" into "portmanteau" morphs throughout the substantival [+N] declensions. This is not universally true: Number and Case appear in "separatist" fashion in a number of Uralic and Altaic languages, most obviously in Turkish, or to a large degree, in the Pama-Nyungan languages of Australia. The second fact, often trivialized, is that number-case affixes invariably appear as suffixes. Yet this too is also a fact specific to Latvian (and European languages in general).

Both facts are captured together in the representation Q, Number-Case. The first fact is shown by the Q bearing the identification "Number-Case." The second fact is captured by the position which Q, Number-Case occupies in the string. The Hallean model requires, indeed, that both facts be inseparable. In other words, certain positions in

[2]Concrete morphemes, conversely, are phonologically invariant, i.e. not subject to allomorphy, except that of readjustment rules. Thus, they have a string instead of a Q. By phonologically invariant it is meant only that their underlying shape is not determined by the morphosyntactic features they carry. Concrete morphemes may undergo readjustment rules (of a feature-changing type) or later regular phonological rules in PF.

the string are identified as bearing certain types of exponence. For languages in which this is invariably true, no problem arises.

Turning once again to the Arabic verbal inflection, I showed in section 0.2.1 that the notions "position" (= Q associated with a string-position), and "exponence type" (= Q associated with specific types of features) cannot always be inextricably linked as they are in the Hallean model. I propose instead that some languages permit Qs (abstract morpheme *positions*) to be underspecified as to the type of features they express.

1.7 Qs and their Labels

Consider once again the Arabic prefix conjugation. We have seen that each verb has at most once prefix and at most one suffix. A plausible representation for this in the Hallean sense would be:

(39) $[[Q_1, ??\ \ [\text{Stem}]]\ \ Q_2, ??]$

The task is to identify what the other half of each Q should be, its label. Consider first the prefixes, given again in (40):

(40) a. Q_1 --> y / ___, 3 f pl
b. Q_1 --> y / ___, 3 m
c. Q_1 --> n / ___, 1 pl
d. Q_1 --> ? / ___, 1
e. Q_1 --> t Elsewhere

The spell-out rules in (40), taken together, rewrite Q_1 as phonological material. As rewriting rules, they are subject to bleeding, so that (40b) bleeds (40c) for example. The ordering (40b) before (40c) follows by the Paninian principle. Because Q cannot appear as a phonological string, rule sets such as (40) also may include an Elsewhere rule, here (40e), which gives *t-* as the default prefix for this paradigm.

Examining (40), we see that, of the different types of features-- Person, Number, and Gender, all three may appear associated with Q_1. Before treating this further, I now review the suffixes as well.

The rules rewriting Q_2 appear in (41):

(41) a. Q_2 --> -ina / ___, 2 f sg
b. Q_2 ---> u / ___, 1 pl
c. Q_2 --> na / ___, f pl
d. Q_2 --> uuna / ___, m pl
e. Q_2 --> aani / ___, dual
f. Q_2 --> -u Elsewhere

As can be seen from (41), all three types of features-- Person, Number and Gender-- appear associated with Q_2 as well as Q_1. As things now stand, the abstract-morpheme treatment is formally identical to the paradigm-model. (To see this, compare (40) and (41) with the rule blocks in section 1.3, (12) and (12B). They are notational variants). We will consider four possible solutions to this puzzle.

The first solution states that BOTH Q_1 and Q_2 have the features of Person, Number, and Gender. Thus the Arabic imperfect verb has the form:

(42) [[[Q_1, Person-Number-Gender [Stem]] Q_2, Person-Number-Gender]

If Person-Number-Gender features reside in a functional head AGR, the movement of the verb to AGR should give an adjunction structure as in (43):

(43) [$_{AGR}$ AGR V] where linear order is undetermined

The template in (42) seems to require that AGR is "in two places at once." The fact that AGR may precede or follow V (since syntactic structure is unlinearized) really doesn't help matters, since AGR must precede or follow V, but cannot do both a the same time in the same form. The AGR abstract morpheme must therefore be copied to the suffix position from the prefix position or vice versa.

Suppose then that copying is permitted. This solution is still vexed by the discontinuous bleeding problem.

Imagine that the Person-Number-Gender features are indeed represented in two places at once. The rules in (40) and (41) operate without complication to yield to the correct forms. Once again, as in the paradigm model, there must be two rules for *-u* (41b) and (41f), and two rules for *y-* (40a) and (40b). The discontinuous bleeding problem has not been solved. Copying essentially permits the wholesale multiple exponence of the

AGR properties, and we have seen that this permissiveness -- principled for the paradigm model -- is too great.

The second possible solution is that Q_1 and Q_2 do have distinct labels. For example, suppose Q_1 is treated as Q_1, Person and Q_2 is in fact Q_2, Number-Gender. The Spell-Out Rules must be modified as in (44) and (45).

(44) a. Q_1 --> y / ___, 3 + Stem + Q_2, pl-f
b. Q_1 --> y / ___, 3 + Stem + Q_2, m
c. Q_1 --> n / ___, 1+ Stem + Q_2, pl
d. Q_1 --> ? / ___, 1
e. Q_1 --> t Elsewhere

(45) a. Q_2 --> iina / Q_1, 2 + Stem + ___, f sg
b. Q_2 --> na / ____, f pl
c. Q_2 ---> u / Q_1, 1 + Stem + ____, pl
d. Q_2 --> uuna / ___, pl
e. Q_2 --> aani / ____, dual
f. Q_2 --> -u Elsewhere

First, observe that the above rules must have nonlocal conditioning environments, that is, in a number of cases, they must be able to "look across" the Stem to a non-adjacent abstract morpheme in order to operate correctly. This violates a constraint originally proposed by Siegel (1977), the Adjacency Constraint.

The rules in (44) and (45) also cannot apply properly owing to ordering paradoxes unless a Q retains its label after it is spelled out. To see this, suppose that the rules in (44) precede those in (45). Since spell-out rules operate cyclically (Halle 1992), either (44) will precede (45) or vice versa. If (44) precedes (45), then (45a) cannot apply properly, since Q_1, 2 will already have been rewritten by (44e) as *t*-. We cannot change the rule to refer to the phonological string *t*- instead of Q_1, 2 since this would derive *t-aktub-iina* for the 3 f sg form. The fact is that Q_2, f sg is rewritten as *-iina* only in forms beginning with a *t*- which is derived from Q_1, 2.

Alternatively, suppose that (45) precedes (44), that is, that Q_2 is on the inner cycle of spell-out rules. A similar problem arises.[3] In fact the same type of paradoxes result if Q_1 is Q_1, Person-Gender and Q_2 is Q_2, Number.

In order to evade these paradoxes, Halle's model must be modified such that the features labelling a Q remain visible even after this Q is spelled-out. But even if this modification is made, the discontinuous bleeding problem is not solved: there are still two rules for *y-* (44a,b) and two rules for *-u* (45c,f).

Another possible maneuver we can make now (maintaining the idea that Q_1 and Q_2 do in fact have invariant labels of some kind) is that Q_1 is Person-Number-Gender together and Q_2 is something else, presumably Tense. Rules (44) and (45) can now be rewritten:

(46) a. Q_1 --> y / ___, 3-pl-f
b. Q_1 --> y / ___, 3-m
c. Q_1 --> n / ___, 1-pl
d. Q_1 --> ? / ___, 1
e. Q_1 --> t Elsewhere

(47) a. Q_2 --> iina / Q_1, 2 f sg + Stem + ___, Imperfect
b. Q_2 --> u / Q_1, 1 pl + Stem + ___
c. Q_2 --> uuna / Q_1, pl + Stem + ___, Imperfect
d. Q_2 --> na / Q_1, f pl + Stem + ___, Imperfect
e. Q_2 --> aani / Q_1, dual + Stem + ___, Imperfect
f. Q_2 --> u / ___, Imperfect

The rules essentially treat all the suffixes as allomorphs of the Imperfect morpheme. This is not entirely implausible: recall from section 1.1 that the subjunctive and jussive conjugations differ by substitution of *-a* and -Ø for imperfect *-u*. Thus, the Tense (or Mood) features clearly condition the suffix allomorphy in some cases, but not always, since *-na* 'f pl' for example, is invariant.

The analysis in (46), (47) however again differs minimally from the paradigm-model analysis because it again fails to explain the discontinuous bleeding phenomena. We are still left with two distinct environments for the spelling-out Q_1 as *y-* in (46a) and (46b).

---

[3](44c), for example, cannot apply properly, since part of its structural description, namely Q_2, pl has already been rewritten as *-u* by (45c) on the previous cycle. We cannot condition this rule with the string *-u* instead of Q_2, pl, since this will entail deriving *n-aktub-u* for the 1 sg form, which also has *-u*. The fact here is that Q_1, 1 is rewritten as *n-* only in forms ending with a *-u* derived from Q_2, pl.

Again, two rules (47b) and (47f) are needed to spell out *-u*. The analysis, just like the paradigm model, commits us to the idea that these are mere accidents. Furthermore, we are again left with quite a collection of complicated allomorphy rules for Imperfect in (47). We have no explanation, for example, why 2 f sg has a special allomorph for Imperfect, for example. This too must be treated as an accident of the morphology.

Additionally, certain parallelisms with substantival morphology are lost, or at best de-emphasized by treating all the suffixes as allomorphs of Imperfect. For example, the allomorph of Imperfect occurring with Q_1, dual, namely *-aani*, is in fact the nominative dual suffix for both adjectives and nouns:

(48) a. rajul-aani kabiir-aani 'two great men'
man- dual.nom great-dual.nom

b. imra?at-aani jamiil-at-aani 'two beautiful women'
woman-dual.nom beautiful-f-dual.nom

(Thackston: 1984:12)

Similarly, the allomorph of the Imperfect appearing with Q_1, pl is the same is the masculine nominative plural suffix (the so-called "sound plural" not formed by prosodic substitution):

(49) mu?min-uuna muxlis-uuna 'sincere believers'
believer- m.pl.nom sincere- m.pl.nom

(Thackston 1984:17)

These facts suggest that *-uuna* and *-aani* in the verbal inflection should be treated as number morphemes, rather than allomorphs of tense conditioned by number features on another morpheme. In other words, Q_2 really ought to be Q_2 Number-Tense. Yet giving Number to Q_2 leads inexorably to ordering paradoxes. Therefore, we are forced to treat the parallelism between the Imperfect allomorphs of the verbal inflection and the Number-Case suffixes of the substantival system as accidents, despite their obvious similarity.

Finally, the simplest treatment of the prefix rules makes the *t-* prefix rule the Elsewhere case as in (46d). Unfortunately this fails to capture the fact that *t-* really represents 2 in some instances and f in others. (See the earlier discussion of the difference between (10) and (11)). The fact that *t* represents 'f' is abundantly illustrated elsewhere in Arabic morphology. For example, the feminine perfect suffix is *-at* in *katab-at* 'she wrote,' dual form *katab-at-aa* 'they (2 f) wrote.' Similarly, the feminine adjectival suffix is also *-at* as in (48b). Treating *t-* as the Elsewhere case as in (46e) is to render this parallelism a synchronic accident. For it could just as well be *b-* or *q-* that would be the elsewhere prefix.

Now let us consider one more possible solution within Halle's 1992 theory. Suppose now that Q_2 is in fact AGR and TENSE and that the prefix position is no more than a stem-augment like the prefix *ge-* in the German past participles such as *ge-fund-en* 'found.'

This solution also fails to capture the discontinuous bleeding effects. Consider the rules which would be necessary:

(50) a. Ø --> y / ___ + Stem, Q_2, 3-pl-f
b. Ø --> y / ___ + Stem + Q_2 3-m
c. Ø --> n / ___ + Stem + Q_2, 1-pl
d. Ø --> ʔ / ___ + Stem + Q_2, 1
e. Ø --> t / ___ + Stem

(51) a. Q --> iina / Stem + ___, 2 f sg
b. Q --> u / Stem + ___, 1 pl
d. Q --> na / Stem + ___, f pl
c. Q --> uuna / Stem + ___, pl
d. Q --> aani / Stem + ___, dual
e. Q --> u / ___, Imperfect

We still need two rules prefixing *y-* to the stem (50a,b) and two rules spelling out -*u* as the INFL morpheme (51b,f). No advantage is gained by treating the prefixes as stem augments. The problem of discontinuous bleeding is still not solved.

1.8 Qs with Variable Labels: Proposal

In this section I introduce my own proposal, which modifies the Hallean view by complicating the mapping from the output of syntax to the input of morphology. In order to capture discontinuous bleeding effects as well as such violations of the Adjacency Constraint as occur in (47), I propose to make Q_1 and Q_2 independent of the location of AGR. As I illustrated in section 0.2.1.1, this entails the splitting of INFL into two locations of exponence.

I envision a mapping as in (52):

(52)	output of syntax:	[[[V] TENSE] AGR]
	Linearization	[[[V] * TENSE] * AGR]
	Merger:	[[V * [TENSE+AGR]]]
	autonomous word-template:	[[Q_1 * stem] * Q_2]

The asterisks represent strict adjacency at the phonological level (Marantz 1988). In the mapping in (52), features in AGR and in Tense combine into a fused M^0 which I will call INFL. Then these features may be linked to either Q_1 or Q_2, as directed by each individual morphological rule.

Each morphological rule accomplishes the following: (1) places a feature or features in a position of exponence (in autonomous word structure); (2) discharges the feature(s); (3) fills the position of exponence (discharges the position) with a string.

In the Arabic case, INFL is required to split into two positions of exponence by the autonomous word structure selected by this INFL. In other words, a [-perfect] INFL is associated with an obligatory prefix and suffix position in the Arabic verb in Phase II of the mapping from Syntax to Phonology. This constitutes the well-formedness conditions on nonperfect verbs (excluding the imperative, which I discuss later).

Spell-out then becomes the choice of the best among available affixes to fulfill the selected template. Filling the template consists in discharging each Q as a string.

Each morphological rule individually places features in positions. For example, the *n-* prefix rule places the features '1 pl' in the prefix position and then discharges them as *n*, filling the Q. Hence each Q receives automatically the label it requires by each individual spell-out rule.

Each rule is associated with an affix which is either a prefix or a suffix as a phonological property. I take this property to be a feature associated with affixes, [±prefix]. The value for this feature I will call the *polarity* of an affix. The polarity of each affix in Arabic dictates which Q its features will be placed in.

In cases where a position is invariably associated with a particular sort of exponence, as in the fused case-number suffixes of Indo-European, then the word template selected by a stem will invariably label the position. For example, one need not learn that each individual case-number affix in Latvian (or other Indo-European) has suffixal polarity: this redundancy is encoded in autonomous morphological structure. In contrast, the polarity of affixes in Arabic must be learned, since INFL splits variously into prefixes and suffixes.[4]

Because morphological rules as I envision them discharge (fill) positions, they continue to give the position-class disjunctions as required. This is entirely analogous to Halle's Q-rewriting, since when a Q is rewritten as a string, the string occupies the position of the Q and two strings cannot occupy a position simultaneously.

More concretely, I will assume the following, ordered morphological rules:

(53)

a.	n-	1 pl
b.	ʔ-	1
c.	t-	2
d.	-aani	dual pl
e.	-na	pl f
f.	-uuna	pl
g.	-iina	f (2)
h.	t-	f
i.	y-	Elsewhere
j.	-u	Elsewhere

discontinuous bleeding — positional bleeding — disjunctive without bleeding

[4]It is probably possible to assume that suffix, i.e. [-prefix] is the default polarity of an affix, and so only the prefixes need be learned with a marked polarity. One reason why suffixing may be unmarked is that this is an aid to parsing: if the stem is normally initial, then word-boundaries also define the left-boundary of the stem in the default instance.

The rules are applied in the order given. The lines show each necessary ordering relationship. Note that some of the ordering relationships apply within positions (to the right of the rules), while others represent disjunctions applying across positions. Of the ones applying within positions, one is not the result of bleeding. We return to this in a moment.

To illustrate what each rule accomplishes, consider the very first rule. In an input with an INFL '1 pl', rule (53a) does the following: (1) the features '1 pl' in INFL are placed in the prefixal position in the autonomous word structure; (2) the features '1' and 'pl' are discharged, so that, for example, rule (53f) will not then apply; and (3) this prefix position is filled by *n-* , so that, for example, rule (53i) will not fill this position by the elsewhere prefix *y-*.

Consider the orderings required in the above rules. All the ones that entail a bleeding relationship can be ordered by the Paninian principle (except (53d) before (53e), which I return to momentarily). These are entirely unproblematic. There is another crucial ordering which does not arise automatically from the Paninian principle: The prefix *t-* '2' (53c) must be ordered before the prefix *t-* 'f' (53h). Otherwise, a 2 f (sg) input would yield *t-aktub-u* rather than *t-aktub-iina*, as shown in the two derivations below:

(54) Correct derivation: (53c) > (53h)
input: 2 f sg
(53c) t-aktub [prefix filled and '2' discharged]
(53g) t-aktub-iina [suffix filled and 'f' discharged]
(53h) cannot apply since prefix position is filled

(55) Incorrect derivation: (53h) > (53c)
input: 2 f sg
(53h) t-aktub [prefix filled and 'f' discharged]
(53c) cannot apply since prefix position is filled
(53g) cannot apply since '2' was not discharged by (53c)
(53j) *t-aktub-u [suffix position filled by elsewhere affix].

In the incorrect derivation, (55), if the *t-* 'f' rule applies first, it will fill the prefix position for a 2 f sg input. Then, *t-* '2' cannot apply, since the prefix position is filled. (53g), the *-iina* rule, requires a discharged '2' to apply. Since (53c) did not apply, '2' will not be discharged, and (53g) will not apply either. The suffix position will then be filled by the default suffix *-u* and ill-formed 2 f sg **t-aktub-u* will result.

There is, however, a principled way to enforce the desired ordering. Observe that all the rules can be aligned according to the principle that person features have greater priority than number features which in turn have greater priority than gender features. This hierarchy is the very same one which in section 0.2.2.2 I argued conditions which feature is deleted when a filter is violated. This convergence of hierarchies is certainly not a necessary fact and therefore provides a surprising result. A considerable portion of the remainder of this thesis will be devoted to testing the hypothesis that the hierarchy which determines affix priority (morphological rule ordering) is in fact the same hierarchy which governs neutralizations.

Combining the hierarchy of features with the principle of Panini gives the following hypothesis regarding rule ordering:

(56) *Spell-Out Ordering Hypothesis*

If a given input can undergo two different spell-out rules the following principles order the rules in the unmarked instance, where one of two situations will obtain:

(1) Panini's Principle: If one rule's structural description is contained in the other's, the rule with the more specific structural description applies first

(2) Feature Hierarchy: If the structural descriptions are disjoint or overlapping, then the rule referring to the hierarchically higher feature applies first

All of the relationships which entail bleeding in (53) fall under condition (1) above, except (53d) before (53e), as I mentioned earlier. The earlier rule has the structural description 'pl dual' and the latter has 'pl f.' This ordering falls under condition (2). The structural descriptions are overlapping, but 'dual' is higher on the hierarchy than 'f' and so 'pl dual' applies first. If the rules apply in the wrong order, a 3 f pl dual form would come out **t-aktub-na* rather than *t-aktub-aani*. The ordering of (53g) before (53h) also follows from condition (2) above, since '2' is higher than 'f' by the hierarchy.

Combining clause (2) of the Spell-Out Ordering Hypothesis with the conception of Impoverishment presented in section 0.2.2, we arrive at the following hypothesis:

(57) *The Feature Hierarchy Hypothesis*

There is a universal hierarchy of morphosyntactic features.
If F and G are morphosyntactic features and F is higher than G on the hierarchy, then:

(1) if *[αF βG] is active at Morphology, then [αF βG] is Impoverished to [αF]

(2) if two spell-out rules, one referring to F, the other to G and not to F, have disjoint or overlapping structural descriptions, then the rule referring to F applies first

This hypothesis is similar to observations made by Kurylowicz in his fifth law of analogy, summarized by Arlotto (1972:141) as:

(58) Kurylowicz's Fifth Law of Analogy

In order to reestablish a central grammatical distinction, a language will abandon one that is more marginal.

In the theory presented here, 'more central' translates to 'expressed by means of higher features on the hierarchy.' As an hypothesis about historical change, Kurylowicz's law, like the Feature Hierarchy Hypothesis, asserts in effect that hierarchically higher properties cannot be regularly neutralized by lower properties, nor will exponents of higher properties ever 'lose' out to lower properties in the competition for expression.

In the next section I turn to the various allomorphs of tense/mood and explain how the present analysis can be extended to incorporate these.

1.9 Explaining Suffix Allomorphy

In this section the jussive and subjunctive moods will be brought under analysis. This leads naturally to a discussion of the relationship between prosodic structure and autonomous morphological structure.

All of the examples that follow are for forms within the inflectional class called the First Binyan. I will assume, following Aronoff (in preparation) that the Binyanim are inflectional class features supplied by derivational rules. These features then select the prosodic shape of the stem and affect its vocalism.

As discussed in section 1.1, the various suffix allomorphs are conditioned by the features of Tense/Mood. The descriptive generalizations are as in (59) and a chart of the relevant forms in the three tense/moods is given in (60):

(59) a. Wherever the imperfect has *-u*, the subjunctive has *-a* and the jussive has no overt ending

b. Wherever the imperfect has a disyllabic suffix (*-uuna, -iina, -aani*), the subjunctive and jussive moods have only the first syllable of this suffix

c. Otherwise, there are no other alternations (*-na* is invariant).

(60)

Imperfect	Subjunctive	Jussive
-u	*-a*	*-Ø*
-uuna	*-uu*	*-uu*
-iina	*-ii*	*-ii*
-aani	*-aa*	*-aa*
-na	*-na*	*-na*

Because the Imperfect, Subjunctive, and Jussive all share a common prosodic Stem shape (i.e. *aktub* < VCCVC or its equivalent in the prosodic terms of McCarthy & Prince 1990a), then these form a natural class in contradistinction to the perfect tense, which has the stem shape CVCVC, e.g. *katab-*.

I propose to express this difference through the feature [± perfect]. The Subjunctive and Jussive are the [-indicative] subtypes of the [-perfect] category, as illustrated in the feature table in (61):

(61) Tense/Mood Categories in Classical Arabic

Name	[±perfect]	[±indicative]	[subjunctive]	[jussive]
Perfect	+	+		
Imperfect	-	+		
Subjunctive	-	-	m	
Jussive	-	-		m

In the table above, [±perfect] is treated as a bivalent feature, whereas [subjunctive] and [jussive] are treated as monovalent features or attributes of Tense/Mood. Owing to the different status of the two feature-types, I use the coefficient [m] (for 'mark') to indicate the presence of a monovalent feature, whereas [+] or [-] indicates the presence of a bivalent feature as well as its value. I postulate monovalent features wherever the absence of the feature (its negative value) appears to play no role.

Returning to the suffix allomorphs in (60) the most striking fact is that the bisyllabic suffixes *-aani, -iina* and *-uuna* appear truncated in the [-indicative] forms, that is, in the subjunctive and jussive. If we do not treat this relationship as one of truncation, but rather one of (suppletive) allomorphy, we are forced to assert (in effect) that it is a mere accident that the [-indicative] affixes are in all cases the first syllables of the [+indicative] affixes. I will not pursue this option.[5]

How then can we express this truncation, i.e. the generalization in (59b)? One way is to ensure that the suffix is maximally monosyllabic in the subjunctive and jussive. In order to express this observation, reconsider for a moment what the suffix is precisely.

Recall that in the Hallean model, each position is occupied by a Q -- a dummy element which stands in for phonological material. What this phonological material is precisely is detailed by rules which spell-out Q. However, nothing in principle prevents us from stating rules which delimit the string which a given Q may appear as. In other words, it is perfectly plausible to state a rule as in (62):

(62) $Q_{Suffix} \rightarrow [_{Q_{Suffix}} \sigma\,] \; / \; \_\_\_\_, \text{[-perfect, -indicative]}$

The rule in (62) expresses the fact that the suffix position is maximally monosyllabic in the non-indicative moods (subjunctive and jussive). Note that rule (62) does not in fact rewrite Q_{Suffix}, in other words, it is NOT structure-changing; rather, it is a structure-building rule which provides the suffixal position with a particular prosodic template. Thus, the features [-perfect, -indicative] are not discharged by rule (62) nor is the suffix position discharged (filled) by the rule.

This rule has the effect of a *morpheme-structure constraint.* No non-indicative verb suffix (of this conjugation) can have a shape larger than a single syllable. One way of understanding such rules is to think of them as redundancy rules operating over the

[5]As Morris Halle has pointed out, a different analysis of these forms would have a special rule supply the suffix *-nV* to all [+indicative] verbs ending in a heavy syllable. The -V- is then disharmonic with the preceding long vowel, and also nonround:

$$V \rightarrow [\alpha hi] \; / \; \underset{[-\alpha hi]}{V} \; C \_\_\_$$

It is of historical interest that all the modern dialects have lost the extra *-nV* which may have been borrowed from the nominal desinences. This historical development is expected if this suffix were added by a costly, special rule such as that above.
Furthermore, the truncation of *-nV* in verbs is curiously parallel to its truncation in nouns when in the construct state. For this reason, this rule may be conditioned by syntactic constituency rather than by morphology per se.

suffixes in this conjugation. Insofar as the affixes realizing a particular position share a prosodic shape, this information can be encoded by means of this kind of redundancy rule.

This kind of prosodic morphology is entirely typical of the general shaping of roots in Arabic, as has been amply demonstrated (McCarthy 1979, McCarthy & Prince 1990a,b). We now extend this shaping to affixes as well. The truncation of suffixes in (59) is on a par with this strategy of prosodic shaping. In other words, the suffixes *-uuna, -iina,* and *-aani* are not in fact truncated in the subjunctive and jussive moods, but rather their second syllables cannot be prosodically licensed (in the sense of Itô 1986), owing to the imposition of the monosyllabic form for Q, Suffix by rule (62).

Introducing rules of form (62) is a considerable enrichment of the theory of morphological realization. In effect, what I am claiming is that not only may the content of an affix reflect its morphosyntactic features, but so also may the *form* of an affix. The most typical structure-building operations on morphological positions are of the prosodic type, i.e. manipulations of syllable structure. Truncations of the type seen in the Arabic suffixes are a subclass of prosodic operations, inasmuch as they establish a maximal prosodic quantity for the affix.

The idea that word positions, or more precisely the strings which come to occupy them, can be prosodically delimited, receives support from another observation from the very beginning of this chapter, namely that all the prefixes of the imperfect conjugation are single segments. Recall that both the paradigm-model (including EWP) and the lexical affix-model have no way of expressing this fact. Why in fact should it be so? One obvious explanation is that the prosodic shape chosen by the [-perfect] verb stem is VCCVC, which lacks an onset.

Onsets are obligatory in all full (i.e. non-functional) words in Arabic. This still does not explain why, for example, there is no verb of the form *CVC-*aktub-*. Such a form would fulfill the onset requirement as well as have an extra prefixal syllable. Nor, really, does it explain why the onset is not simply a phonological default onset such as *w-* or ʔ-. In fact, the onset is ʔ- in the imperative ʔ-uktub-Ø (m sg), ʔ-uktub-ii (f sg), ʔ-uktub-aa (dual), ʔ-uktub-uu (m pl) and ʔ-uktub-na (f pl). The imperative INFL does not select an obligatory prefix position in autonomous word-structure, but the *phonology* of Arabic still requires an onset, and so default ʔ- appears in all the imperatives. This establishes that the morphological requirement of a prefix cannot be reducible to the phonological requirement of an onset: the latter can be met independently of the former.

In the non-imperatives, the onset is therefore an affix with a position in autonomous word structure, not just a phonologically required onset. To express this, I propose a rule which delimits the prefix to a single segment:

(63) Q_{Prefix} --> $[_{QPrefix}$ X]

This rule too takes the form of a redundancy rule operating over the affixes realizing the prefix position. One advantage to such a rule is that it allows certain similarities between prefixes and suffixes which appear in the perfect. For example, *n-* '1 pl' appears as a suffix, *-naa,* in the perfect tense. If the affix is really *naa* but is in effect "pre-truncated" by rule (63), then only one '1 pl' affix need be learned. Similarly, *t-* '2' appears as *-ta* in the perfect and *t-* feminine as *-at* in the perfect. Rule (63) ensures that these affixes have only one segment (consonantal, since it must be an onset) when appearing as prefixes.

The suffixal position is not specified for prosodic shape, so it can be left either undefined (i.e. autotemplatic) or its shape can be dynamically delimited by a rule such as (62).

Now let us consider the relationship which autonomous word-templates as in (64) below have with prosodic templates. We know, for example, that the Arabic [-perfect] stem has the form VCCVC (or its moraic equivalent; I use CV-skeleta for simplicity here). We also know that the imperfect verb has a specific morphological template, and that the prefix position is limited to a single segment. The relation between these is shown in (64):

(64)	prosodic template	C	VCCVC
		|	|
	word-template	[[Q_{Prefix} *	Verb] * Q_{Suffix}]

A way to generate the structure in (64) is to say that a [-perfect] V+INFL stem selects a particular prosodic template (two heavy syllables with no onset) and a word-template (a prefix and a suffix position: Q_{Prefix} and Q_{Suffix}).

The particular prosodic shape chosen (two heavy syllables without onset) is a *stem formative* of the verb radical. The fact that the stem formative happens to be a prosodic shape rather than an affix, as in the themes of Latvian or Huave or the word markers of Spanish, is, while not particularly obvious, not surprising, once autosegmental phonology is permitted. This was essentially McCarthy's original observation.

Just as each noun in Latvian nouns requires a Theme, each Arabic verb requires a prosodic shape, and this requirement is clearly a language-specific matter of autonomous word structure.[6] Furthermore, the vocalism of the verb is also a stem formative. In some cases, the vocalism is predictable, in others, selected by the radical. For example, the first vowel of the [-perfect] forms is *a* but the second may be any of the three vowels *a i u*: *y-aktub-u* 'he writes,' but *ya-ʕlam-u* 'he knows, learns,' and *y-anzil-u* 'he descends.'

These requirements can be diagrammed as below:

(65) Q, Vocalism
|
Q, Prosody
|
Verb Radical

The apparent peculiarity of such nonconcatenative morphology derives from the superimposition of the stem formatives by Tier Conflation (McCarthy 1979). But otherwise, the Arabic system is not any different from the Latvian one: while the Prosody and the Vocalism of a verb may show allomorphy depending on tense/mood, tense/mood may be realized directly in the suffix position.

Given the truncation effect that rule (62) simulates, the only allomorphy among the suffixes in (61) remaining to be explained is that of the elsewhere cases *-u*, *-a* and -Ø. To this end, I replace the one Elsewhere rule of the last section with the three rules in (66):

(66)	a. -perfect, +indicative	-u
	b. jussive	-Ø
	c. Elsewhere	-a

The subjunctive ending *-a* is treated here as the elsewhere case on analogy to the unmarked 3 m sg suffix *-a* in the perfect *katab-a* 'he wrote.'

Now I review more explicitly the derivation of a [-perfect] verb. Let us take as an example *t-aktub-ii* 'you (f, sg) write (subjunctive).' Assume that after Rebracketing, the output of syntax is:

---

[6]Just as not all Indo-European nominals require themes, not all Semitic verbs have prosodic skeleta. McCarthy (1991) has argued that in Akkadian, certain verbs, while appearing to have the prosodic skeleta familiar from Arabic, in fact arise "autotemplatically," i.e. solely in virtue of epenthesis and other phonotactic rules of the language. In the terms presented here, these verbs have no prosodic stem formative. See also section 3.1.2 on Free Placing of affixes.

(67) $[[_{V}\ \mathit{k\,t\,b}\,]\ [_{INFL}\ 2, \mathrm{f},\ \text{-perfect},\ \text{-indicative}]]$

Because Arabic verbs must conform to the autonomous word structure in (65), the features in (67) will be mapped onto this structure. Moreover, as a property of autonomous word structure, nonperfect verbs in Arabic also select the morphological structure:

(68) Q_{Prefix} * Verb * Q_{Suffix}

There must be a rule which identifies the Prosody of the verb:

(69)

```
                          σ
                         / \
                    R   /  R
                    /\  / / \
Q, Prosody -->      VC  C VC
        |
[binyan 1] + [-perfect]
```

Second, the Vocalism *a u* is derived by

(70)

```
Q, Vocalism --> a u  /   ______
                           |
                        VCVCVC
                      [binyan 1 ...]
```

These two rules and the two conditions on well-formedness give this output:

(71)

```
              a    u
              |    |
   [QProsody VCCVC]
              | |  |
   QPrefix*  k t   b  * QSuffix
              ↑
       [[V k t b ]  [INFL 2, f, -perfect, -indicative]]
```

The Verb root position in autonomous word structure is automatically filled with the radical *k t b*.[7] I notate this discharge by an arrow ↑. The rules which prosodically delimit the affixes (62) and (63) now apply, giving:

(72)

```
              a   u
              |   |
      C    [VCCVC] [σ(C)V(X)]
      |      | | |       |
  QPrefix*  kt   b  * QSuffix
              ↑
      [[V k t b ]  [INFL 2, f,  -perfect,  -indicative]]
```

All that remains now is the splitting of INFL into the subject and prefix positions. The following rules apply:

(73) a. t- 2
b. -iina f (2)

The second syllable of the suffix cannot link to the prosodic template, and therefore deletes by Stray Erasure:

(74)

```
              a   u
              |   |
      C    [VCCVC] [σ(C)V(X)]
      |      | | |       | |
      t    * kt  b  *    i i    n a
              ↑
      [[V k t b ]  [INFL 2, f,  -perfect,  -indicative]]
```

The arrows show the linking of features to positions, which are realized (identified) with phonological material.

Tier conflation applies giving the phonological form *t-aktub-ii*. Syntactic structure (the bottom line in (74)) is no longer licensed, and the well-formed verb enters Phonology.

---

[7] I will assume that by convention the root position in autonomous word structure is automatically identified with the phonological material identifying the root in syntax.

1.10 Hierarchy among the AGR and T Suffixes

Consider now the full set of feature-discharging rules for the [-perfect] inflectional affixes:

(75)

a.	n-	1 pl
b.	ʔ-	1
c.	t-	2
d.	-aani	dual pl
e.	-na	pl f
f.	-uuna	pl
g.	-iina	f (2)
h.	t-	f
i.	-u	-perfect, +indicative
j.	-Ø	jussive
k.	y-	Elsewhere
L.	-a	Elsewhere

Several new crucial orderings are required and are noted in the above array. These orderings express the fact that the tense/mood suffixes *-u* and *-Ø* appear only if the suffix position is not filled by an affix discharging AGR features. For example, the 3rd person masculine plural form in the jussive is not **y-aktub-Ø* but rather *y-aktub-uu*. Therefore, the rule identifying the suffix with the feature pl (75f) must apply before the jussive rule (75j).

In all these orderings, tense/mood features are the very lowest in terms of priority. This does not follow from the hierarchy which motivates neutralization, nor from any syntactic principles. Since T is closer to the verb than AGR at the output of syntax before merger, then if the pre-merger structure is used to motivate rule ordering, we should expect T features to be visible first. On the other hand, because merger does apply, T and AGR become a single fused M^0, and therefore T is not then closer to V.

One possibility is that in fact T has fused with the verb rather than with AGR. In this analysis, the tense features become like inherent properties of the verb stem, and therefore are able to condition allomorphy of any peripheral affixes. This view allows the suffixes *-u*, *-a* and -Ø to be various allomorphs of the Elsewhere affix, as conditioned by the tense features on the verb:[8]

(76)	a.	-u	Elsewhere (-perfect +indicative)
	b.	-Ø	Elsewhere (jussive)
	c.	-a	Elsewhere

It should be carefully noted that such multiple elsewhere allomorphs (like all allomorphy alternations) can be conditioned only by inherent features of the stem or by previously discharged features. It is predicted, then, that allomorphy -- indeed all sorts of allomorphy -- can be sensitive only to inherent properties of the stem or to properties realized more closely to the stem. Thus, we derive the effects of Carstairs' (1987:193) Peripherality Constraint:

(77) Peripherality Constraint

> The realization of a property ... may be sensitive inwards, i.e. to a property realised more centrally in the word-form (that is, closer in linear sequence to the root), but not outwards to an individual property realized more peripherally (further from the root).

Whether T merges with the verb or with AGR in classical Arabic is unresolved here, pending a more extensive investigation of the stem allomorphy of the various verb tenses.

1.11 Zero Affixes

One of the most persistent debates in the study of morphology centers on the question of Ø-morphs (Matthews 1974 has good discussion). Among recent work, opinions range from staunch support, as in Pesetsky's (1991) Zero-Syntax, to dismissal of zero-morphs entirely (Zwicky & Pullum to appear).

Since the advent of the theory of empty categories, opinion has converged on the idea that null elements play a role in syntax. That a particular X^0 is phonologically empty is not controversial.

Insofar as affixes and X^0s are isomorphic, it would appear that the existence of null X^0s demands the existence of null affixes as well. Yet I have shown that affixes are not

[8] I wish to thank Morris Halle for pointing out this possibility to me.

isomorphic to X^0s, so this argument no longer holds quite as directly. We must establish an explicit definition of zero-morphs at the *morphological* level.

In this thesis, zero-morphs have a particular interpretation as follows. Observe that the Arabic jussive suffix is -Ø. What is the real content of asserting that the suffix is nothing, a null?

For Halle (1992), a zero-suffix results when a Q is rewritten as Ø. Because Qs are not well-formed entities at Phonological Form, every Q must be replaced, and in cases where a postulated position has no string associated with it for a given word, Halle's theory demands a rule replacing Q by Ø.

In the theory presented here, I have defined Qs as obligatory positions of exponence which may figure in autonomous word structure. Insofar as these positions are obligatory, no prefix-conjugation verb in Arabic will be well-formed without a suffix. Since the suffix -Ø is associated with a particular property, namely the jussive mood, which has no other principal exponent, we are justified in postulating a *contentful* -Ø affix.

The identity of the jussive forms lies precisely in the fact that they are suffix-less. This is encoded by the special rule (75j) conditioned by the feature [jussive]. Since jussive is a morphosyntactically marked category, the system shows *iconicity* if a more marked category morphosyntactically is more marked morphologically. Yet a zero-suffix would appear, prima facie, to be precisely the opposite: maximally unmarked. On the contrary, in the system as analyzed here, the zero-suffix is in fact marked *with respect to the system as a whole*. Hence there is an iconic relationship as expected. If we abandon the idea that zero-suffixes exist (or, equivalently, that there is an obligatory position which is filled by special rule), then the iconicity of the system is lost.

Not all zeroes have content, however. I will be arguing in chapter 3 that a contentful Ø may occur only where a position is obligatory in autonomous morphological structure. In Berber (section 1.18) and many languages with multiple argument agreement, affixes need not be licensed by positions-of-exponence in autonomous morphological structure. In such languages, since no position is obligatory, there can be no contentful zero-affixes. Contentful zero-affixes like the Arabic -Ø (and also like the enclitic -Ø in Mam, to be discussed in section 2.1.6), however, discharge features and fill obligatory positions.

1.12 Historical Reanalysis and Discontinuous Bleeding

The imperfect conjugation in Arabic is one representative of what is known as the prefix conjugation within the Afroasiatic superfamily of languages. Because of its extreme antiquity, the prefix-conjugation can be traced with relative ease from its earliest attestations in Old Akkadian c. 2500 B.C.E. to many of the modern languages of the Near East and northern Africa. The historical changes which develop within this conjugational pattern form the focus of the next several sections, and provide a useful arena in which to test any analysis. I will be arguing that these historical changes can be analyzed most simply in the theoretical framework advanced in the previous sections for the Arabic prefix conjugation. In particular, I will argue that certain languages gained or acquired filters (i.e. cooccurrence restrictions for morphological features), lost certain affixes, or reanalyzed the content of affixes. I endeavor to show that the systemic adjustments that accompanied these changes follow automatically, and certain "curious" problems of historical Afroasiatic morphology receive new and natural explanation.

1.13 The Fate of *-uuna* and Gender in the Plural

Recall from section 1.5 that the affix-based model and the word-paradigm model are both capable of generating the correct forms in the Arabic prefix conjugation. The arguments I have presented in favor of a new analysis of the facts do not hinge on descriptive adequacy. But to express the facts, both the affix- and paradigm-based analyses have to incorporate special rules or complicate the feature content of particular affixes. In this section, 1.13, I consider the first case of discontinuous bleeding -- that of *-uuna* 'pl' by *n-* '1 pl.' Historical developments which I now detail can be expressed easily and simply if *-uuna* is 'plural' alone and not further specified by either (1) a feature [-1], which might exclude its appearance in the 1 pl, or (2) 'masculine,' which, by appeal to the lack of gender in the 1st person, might likewise exclude it from 1 pl. These developments support my contention that *-uuna* is plural alone and that discontinuous bleeding should not be captured by extra features but by the workings of the theory itself.

1.13.1 Gender in the Plural

A number of modern Arabic dialects have lost gender distinctions in plural, for example, Egyptian Arabic (Aboul-Fetouh 1969), the dialects of Bahrain (Al-Tajir 1982) and Abu Dhabi (Qafisheh 1977), as well as all the dialects of the Maghreb, including

Libya, Tunisia, Algeria, and Morocco (Marçais 1977). A simple development from the classical pattern may be seen in Egyptian Arabic:

(78) The Egyptian Arabic Prefix Conjugation (Imperfect)

yi-ktib	3 m sg
ti-ktib	3 f sg
ti-ktib	2 m sg
ti-ktib -i	2 f sg
?a-ktib	1 sg
yi-ktib-u	3 pl
ti-ktib-u	2 pl
ni-ktib	1 pl

Among the changes observed in this dialect are the change of *-uuna* > *-u*, *-iina* > *-i*, and *-u* > Ø, and the ablaut of the initial *a* to *i* in all but the 1st person singular.[9]

To express the Impoverishment of gender in the plural, I propose that those dialects of Arabic which lost gender in the plural acquired the filter *[pl f]. Following the Feature Hierarchy Hypothesis, the gender feature f, being hierarchically subordinate to the number feature pl, was lost. We can express this graphically by a delinking rule:

(79)
$$\begin{array}{c} \text{pl} \\ \not| \\ \text{f} \end{array}$$

The application of (79) makes all plurals be unspecified for gender (recall that f is privative in my theory, and so masculine is unmarked).

The filter in (79) must be stipulated to apply only to verbs and pronouns. Nouns retain their inherent gender (m or f) and adjectives continue to inflect for gender and number, but verbs and pronouns cease to show any gender distinction in the plural. For example, in (80) are tabulated the 3rd person pronouns of Classical and Egyptian Arabic (the latter from Aboul-Fetouh 1969:85):

(80)

	Classical	Egyptian
3 m sg	huuwa	huwwa
3 f sg	hiiya	hiyya
3 m pl	hum	humma
3 f pl	hunna	humma

[9] By analogy with the third person pronominal *hum* the plural ending can appear as *-um* rather than *-u*

Observe that there is now no way to treat the Egyptian Arabic plural suffix *-u* as masculine plural, since the feminine plurals in fact also take this suffix in Egyptian. Recall that this was one possible maneuver out of the problem posed by the discontinuous bleeding of *-uuna* 'pl' by *n-* '1 pl,' both for paradigm models and affix-based models.

In the modern Arabic dialects which have lost gender in the plural there seems no way of avoiding treating *-u* < *-uuna* as simply the suffix for plural *alone*. We cannot block affixing it to a 1 pl form on the basis of a gender feature. Recall that EWP requires a special rule of some sort to prevent this from happening. For classical Arabic, I showed that EWP might have recourse to a disjunction of the type

(81) a. V, 1 pl --> V+u
b. V, pl --> V+uuna
c. Elsewhere, V --> V+u

For Modern Egyptian, the same set of rules would have to be:

(82) a. V, 1 pl --> V
b. V, pl --> V+u

What is noteworthy about (82) is that, in order to prevent the 1 pl form from being suffixed by rule (82b), there must be a previous rule disjunctive with this rule, and this previous rule must do nothing, i.e. rewrite a string as itself. Alternative strategies, such as permitting certain disjunctions across rule blocks, lead to ordering paradoxes, as I showed earlier.

Such identity rules are necessary only in a theory which in principle cannot express the discontinuous bleeding of *-u* 'pl' by *n-* '1 pl.' In a theory in which discontinuous bleeding can be directly expressed, identity rules are rendered superfluous.

It is certainly not the case that the discontinuous bleeding of the plural suffix by 1 pl prefix is an historical accident. This can be verified directly by examining the identical pattern in another Afroasiatic language of the Northern Cushitic branch, Beja (also known as Hadareb, Bedauye, Bedawie, Bedawye, etc.), spoken in Sudan and Northern Ethiopia. The following paradigm is from Zaborski (1974:12) citing Hudson (1974:133):

(83) Prefix Conjugation in Beja *l w* 'to burn'

ʔi-líw	3 m sg	'he burned' etc.
ti-líw	3 f sg	
ti-liw-à	2 m sg	
ti-liw-ì	2 f sg	
a-líw	1 sg	
ʔi-liw-nà	3 pl	
ti-liw-nà	2 pl	
ni-líw	1 pl	

The remarkable similarity between the Cushitic prefix-conjugation, as in Beja, and the Semitic pattern, as in Arabic, was a decisive factor in postulating the Afroasiatic family of languages as early as Meinhof's study (1912:138).[10] A similar pattern of verbal prefixes appears in the East Cushitic languages ʕAfar and Saho, spoken in Eastern Ethiopia and Djibouti. Three paradigms of Saho are presented below, from Zaborski (1974:22ff) citing Conti Rossi (1913:164):

(84) Saho Prefix Conjugation *g d f* 'to kill'

	present	past
3 m	y-agdif-é	y-igdif-é
3 f	t-agdif-é	t-igdif-é
2	t-agdif-é	t-igdif-é
1	agdif-é	igdif-é
3 pl	y-agdif-ín	y-igdif-ín
2 pl	t-agdif-ín	t-igdif-ín
1 pl	n-agdif-é	n-igdif-é

[10]Afroasiatic includes Semitic, Cushitic, Omotic, Chadic, Berber, and ancient Egyptian. These branches are not necessarily equally close. Reflexes of the prefix conjugation appear in Cushitic, Semitic, and Berber. Whether certain preverbal clitics of Chadic are indeed reflexes of the prefix conjugation, as Meinhof (1912) originally surmised on the evidence of Hausa, has been challenged, for example, by Mukarovsky (1983). Omotic lacks the prefix conjugation entirely; in fact, Omotic verbs show little if any inflection (Fleming & Bender 1976), and ancient Egyptian shows at least no overt trace of the prefix-conjugation. Whether the prefix-conjugation is extremely archaic (Klingenheben 1956), as is the traditional view, or is diagnostic of a subfamily within Afroasiatic, or is an independent development in the various branches (unlikely on my view), remains a topic of debate among Afroasiatic scholars (see Hodge 1971:17-19).

Here we see the same type of system as in classical Arabic: the default mood suffix *-é* appears when *-ín* 'pl' does not already fill the position.[11] Again, there is the same discontinuous bleeding of the plural suffix by the '1 pl' prefix.

On the basis of the evidence from the more archaic representatives of the prefix conjugation in Cushitic, we can safely assume that the discontinuous bleeding relationship between the pl suffix and the 1 pl prefix is a remarkably stable systematic property of Afroasiatic as a whole, antedating the division of Semitic from Afroasiatic: it is both highly archaic and highly resistant to change. It is neither the result of accidental homophonies nor of separate identity rules. Even more compelling evidence for discontinuous bleeding comes from the reflex of the prefix-conjugation in Berber, but this will be discussed in section 1.18.

1.13.2 The Plural Suffix as Specifically non-1st Person

Another possible solution to the discontinuous bleeding problem is to treat *-u* as a specifically [-1 pl] suffix, which requires introducing [-1] as a potential value into morphological representations (a less constrained theory of features).

Suppose, for the moment then, that *-u* is in fact treated as [-1 pl], in other words, a plural suffix specifically unable to appear in 1st person forms. If this move is made, we have no natural way to explain another historical development which I now detail.

In the Maghrebi dialects of modern Arabic generally, the 1st person singular ?- is replaced by the reflex of classical *n-* , that is to say that, the 1 pl prefix becomes the generalized 1st person prefix, as shown in the following paradigm from the dialect of Arabic spoken by Tunisian Jews (Cohen 1975:94):

(85) The Tunisian Jewish Arabic Imperfect

y-ə́ktəb	3 m sg
t-ə́ktəb	3 f sg
t-ə́ktəb	2 sg
n-ə́ktəb	1 sing
y-ə́ktb-u	3 pl
t-ə́ktb-u	2 pl
n-ə́ktb-**u**	1 pl

[11]The suffix *-é* is presumably an exponent of mood, since the subjunctive has *-ò* instead. Cf. the substitution of *-a* for *-u* in classical Arabic.

As Cohen writes (1975:96):

> Without a doubt, the weakening of *hamza* [glottal stop] must have played a role in the extension of *n-* to the 1st person singular from the 1st person plural. This extension, and the generalization to the plural of the desinence *-u* of the 2nd and 3rd persons are connected facts for which it is not possible to determine the order of chronological succession nor orient the causal relationship [trans. mine].

Given the hypothesis that *-u* represents plural alone here, we can in fact make a prediction about the chronological succession. When ?- was lost and *n-* was reanalyzed as '1' instead of '1 pl', then discontinuous bleeding was no longer motivated, and the plural suffix *-u* was applied to all the plural forms.[12] We have no need to countenance a change in the suffix itself, since, as the elsewhere suffix for plurals, it naturally stepped in when not constrained from doing so by discontinuous bleeding of pl to the prefix position by *n-* as 1 pl. This change is shown below:

(86)	Classical			Pre-Maghrebi (cf. Egyptian)			Maghrebi Arabic	
	1 pl	n-		1 pl	n-	>	1	n-
	1	?-		1	?-	>	Ø	
	pl	-uuna	>	pl	-u		pl	-u

If, on the other hand, *-uuna* > *-u* is a specifically [-1] suffix, its extension to the 1 pl form is not automatic at all; in fact, it should be inhibited. In order for *-u* to be extended, it must lose its [-1] feature. There must be two reanalyses: *n-* 1 pl > 1, and *-u* [-1 pl] > pl. Positing two reanalyses cannot then explain the causal relationship between the loss of ?- and the extension of *-u*. For the [-1] feature of *-uuna* > *-u* could have simply been lost independently. Yet this never occurs, not in ANY attested reflex of the Afroasiatic prefix conjugation. Only when *n-* is reanalyzed as '1' in the Maghreb dialects does *-u* extend to 1 pl.

On the other hand, the extension of *-u* to all the persons in the Maghreb dialects follows naturally in the model in which *-u* is plural alone, and its failure to appear in the 1 pl is captured by discontinuous bleeding and not a special feature [-1].

[12]Bergsträsser (1928) makes an entirely inverse speculation. He suggests that *-u* was extended to the 1st person plural forms and that this then "permitted" the extension of *n-* to the 1st person singular forms. As Cohen points out, however, there is no attestation of this change, so we cannot appeal to any facts to settle the question definitively.

To summarize, historical facts strongly suggest that discontinuous bleeding as I have analyzed it is a linguistic reality, both quite archaic and quite stable. The loss of gender in the plural in a number of the modern Arabic dialects, along with the extension of plural *-u* to the 1st person only where there was concomitant loss of *n-* suggest strongly that this case of discontinuous bleeding is no accident, and any correct morphological theory must be able to express it.

1.14 Loss of a Filter: 1st Person Duals

I showed in section 1.11.1 that Egyptian Arabic gained the filter *[pl f] for verbs and pronouns, and that, in accordance with the Feature Hierarchy Hypothesis, the feature 'f' was deleted in all such combinations at Morphology. Other things being equal, we should also expect to find a historical change in which a filter is *lost*. In this section I show that precisely this did happen in the development of 1st person dual forms.

Recall that Classical Arabic has the filter *[1 dual]. The dual number is never distinguished for the 1st person, although such forms as *ʔ-aktub-aani '1-write-dual' could be generated by an input '1 pl' AGR. To evade this, I have argued that a filter *[1 dual] is automatically assumed by the child learning the grammar until he or she finds positive evidence to switch the filter off. The filter impoverishes a '1 dual' AGR, neutralizing it with a '1 pl' AGR.

Recall from 0.2.2.4 that duals are unusual in that they may be optionally enhanced by an otiose 'pl' (i.e. [-sg]) specification, or this enhancing 'pl' may be absent morphologically. If a 1st person dual category were to arise, we should expect two options: either it is enhanced by 'pl' -- formed by analogy to 1 pl -- or it is not enhanced, i.e. analogized from the 1 sg form.

Both options are represented in Semitic, the former by Ugaritic and the latter by the South Arabian languages Meħri and Soqoṯri.

The South Arabian languages Meħri and Šxawri are spoken by small populations in the southern region of the Arabian peninsula (Yemen and Oman), while Soqoṯri is spoken on the island of Soqoṯra, south of the Arabian peninsula (part of Yemen). The prefix conjugation for Meħri (Johnstone 1987) is shown below:

(87) The Meħri Imperfect: *r k z* 'to straighten'

	sg	dual	plural
3 m	yə-rūkəz	yə-rəkz-ō	yə-rəkz-əm
3 f	tə-rūkəz	tə-rəkz-ō	tə-rəkz -ən[13]
2 m	tə-rūkəz	tə-rəkz-ō	tə-rəkz -əm
2 f	tə-rēkəz (-i)[14]	tə-rəkz-ō	tə-rəkz -ən
1	ə-rūkəz	**ə-rəkz-ō**	nə-rūkəz

Observe that the 1st person dual form ə-rəkz-ō has the prefix ə- '1' and the suffix -ō 'dual,' and is thus analogous to the ill-formed *ʔ-aktub-aani '1-write-dual' of Classical Arabic. My analysis is committed to the idea that only the filter prevents this form from being produced in Classical Arabic, which I take to be approximately the Proto-Semitic system, following standard assumptions in Brockelmann (1961) or Moscati (1964). When the Proto-Semitic *[1 dual] filter is switched off, as in Meħri, then the expected analogous forms appear automatically.[15]

A similar pattern may be seen in Soqoṯri (Bittner 1918:73).[16] The dual forms cognate with the Meħri forms are shown below:

(88) The Soqoṯri prefix-conjugation dual *k t b* 'write'

3 m	i-kéteb-o
3 f	te-kéteb-o
2 m	te-kéteb-o
2 f	te-kéteb-o
1	e-kéteb-o

[13]The appearance of *t-* here is discussed in the next section.

[14] The suffix -i > Ø in some forms.

[15]Actually, this form is generable only when the category 'dual' is not enhanced by plural. Otherwise, we should expect *n-aktub-aani* with an enhanced plural prefix. I return to this matter shortly.

[16]Interestingly, Bittner recorded no duals in Meħri or Šxawri, as can be discerned by his remark (p. 73), "Neben dem Singular und dem Plural hat das Soqoṯri *gegen* das Mehri (und auch *gegen* das Šẖauri) auch einen Dual, und zwar für alle drei Personen, also auch für die 1. Person. [itals. mine]" Whether duals -- and in particular 1st person duals -- were absent from Bittner's data or in fact absent from the language Meħri in the 1910s when his data were collected in unclear to me. Obviously, by the 1960s and 1970s when Johnstone worked on Meħri, the dual was fully productive, including for the 1st person.

It should be apparent that any theory which does not incorporate filters will have trouble expressing this change naturally. For example, if the dual suffix is subcategorized to appear only in [-1] categories in Arabic, then one could appeal to loss of this subcategorization in Meħri, for example. But this would make the reanalysis a property of the individual affix and not of the system as a whole. This is clearly incorrect, since Meħri has also innovated a 1 dual category in the suffix (perfect) conjugation:

(89) The Meħri Suffix Conjugation *r k z̧* 'to straighten'

	sg	dual	plural
3 m	rəkūz	rəkəz-ō	rəkáwz
3 f	rəkəz-ūt	rəkəz-t-ō	rəkūz
2 m	rəkə́z-k	rəkə́z-k-i	rəkə́z-k-əm
2 f	rəkəz-š	rəkə́z-k-i	rəkə́z-k-ən
1	rəkə́z-k	**rəkə́z-k-i**	rəkūz-ən

There are several points of interest about the above paradigm.

The first is that 1st person dual rəkə́z-k-i is formed from the 1 sg form rəkə́z-k and not from the 1 pl form rəkūz-ən. Thus dual is not enhanced by the plural in this conjugation either. The 1 dual forms in both conjugations are formed synthetically, that is to say, solely from affixes which are pre-existent in the system. In the prefixing conjugation the two affixes are ə- '1' and -ō 'dual', in the suffixing conjugation -k '1' and -i 'dual.' In order to prevent analogous forms in Classical Arabic, we must appeal to a filter: when the filter is switched off, as in Meħri, these forms are automatically generated. This provides the evidence that historical change may consist only in the loss of a filter, with new forms automatically being built up with the available affixes.

The dual suffix is *-i* in the 1st and 2nd persons but -ō in the 3rd person. It appears that *-i* is cognate with the Arabic nominal dual ending in the oblique case *-ay*, whereas -ō is cognate with dual *-aa(ni)*, which is the verbal dual ending as well as direct case nominal dual.[17]

Given that there exists a dual suffix *-i* and the fact that duals appear to be formed off of singulars in this instance, we should expect a 2 f du form *rəkəz-š-i, with the 2 f

[17]The *-i* dual suffix can be seen in the 3rd person in Sabaean, an ancient South Arabian language known only from unvowelled inscriptions. Brockelmann (1961:576) gives two examples:

(a) ḥħðθ-y 'they two renewed: sie beiden haben erneuert'

(b) šmt-y 'they two set up: sie beide haben aufgestellt'

The origin of the distribution of the dual *-i* vs. -ō in Meħri remains open.

suffix -š. In fact, the 2 f dual form is homophonous with the 2 m dual form: both are rəkəz-k-i, formed from the 2 m sg form rəkəz-k.

This result is predicted by my analysis since 2 f dual is impoverished by the filter *[2 f dual], as in Classical Arabic. When in a [2 f dual] AGR the 'f' is deleted (as a result of the filter and the hierarchy, as discussed earlier), the result is [2 dual], i.e. the unmarked masculine form:

(90)	Input	2 f dual	
	Filter	*[2 f dual]	
	Impoverishment	2 dual	
	Rules	2	-k
		dual (1,2)	-i

The form rəkəz-k-i provides evidence that Impoverishment actually applies to the input [2 f dual], preventing the bad form *rəkəz-š-i which could be generated from the fully specified input.[18]

The only other 1 dual forms in Semitic appear in a few examples from Ugaritic, a North-West Semitic language preserved only in unvowelled texts from c. 1300-1100 B.C.E. Gordon (1947a:59, §9.4) gives the following examples:

(91) a. ql-ny 'we (dual) have prostrated ourselves'
b. [m] ǵ-ny 'we (dual) came'

The 1 dual suffix is attested as unvowelled -ny, which is reconstructed as *-na-yā. The same suffix appears also as a pronominal clitic in possessive construction (92b) and with a preposition (92a):

(92) a. adt-ny 'our (dual) lady'
b. ʕm-ny 'with us (dual)' (Gordon 1947a:28 §6.21)

The 1 dual suffix appears to be formed from the 1 pl *-na* plus dual suffix -(y)ā. It is thus a case in which a dual form is enhanced by the specification 'pl.'

[18]This argument cannot be made for the prefix conjugation since the 'f' feature cannot be spelled out since both the prefix position and suffix position are filled by higher ranking affixes: *t-* '2' and -ō (Ar. *-aani*). If 'f' were not deleted by Impoverishment, it would still fail to appear morphologically in these forms.

Whether Proto-Semitic had 1 dual forms or not is unclear. Some scholars take the 1 dual forms in Ugaritic as suggestive of an archaic 1 dual which was lost in all branches of Semitic other than Ugaritic (later, perhaps, to resurface in South Arabian). In support of this notion, the cognate 1 dual suffix *-ny* of Ancient Egyptian is generally cited.[19] Within the present framework, one cannot know whether Ancient Egyptian and Ugaritic both independently lost the *[1 dual] filter and automatically innovated these forms or whether Semitic (other than Ugaritic) gained the filter after Semitic separated from Ancient Egyptian. Obviously, the existence of 1 dual forms in both Ugaritic and Ancient Egyptian does not in itself prove the archaism of these forms, since all that need happen for these forms to appear is for the filter to be lost.

To summarize, the loss of gender in the plural in Egyptian Arabic shows the case of gaining a morphological filter through historical change. The analogical extension of the dual suffix to the 1st person in Meħri and Soqoṯri shows the opposite: historical change consisting merely in the loss of a filter. Other 1 dual forms, enhanced by 'pl' and formed from the 1 pl form with added dual suffix, appear in Ugaritic. It is unclear whether these forms are derived from Proto-Semitic or not: in any case, either Ugaritic lost the filter *[1 dual] or the rest of Semitic gained it. The point, however, is that all these historical changes, traditionally explained as analogy, can be reinterpreted here merely as loss or gain of filters. Analogical forms are then generated automatically by the pre-existent rules of each system, while forms which are lost cannot be generated since the inputs required to generate them have been simplified by Impoverishment.

[19]Ancient Egyptian and Semitic are usually assumed to be separate branches of Afroasiatic. Forms appearing in Semitic daughter languages with cognates in Ancient Egyptian are safely assumed to be Proto-Semitic (excepting borrowings, of course).

1.15 The *y-* and *t-* Prefixes as Defaults

It is usually assumed that the distribution of *y-* and *t-* in the prefix conjugation of classical Arabic is the same as that of Proto-Semitic (Brockelmann 1961 §6.20), although Akkadian presents a complication to be discussed shortly. Recall that *t-* appears in all 2nd person forms, and 3rd person feminine *except* 3 f pl, as can be seen in the 2nd and 3rd person forms retabulated below:

(93)

	singular	dual	plural	
	y-aktub-u	y-aktub-aani	y-aktub-uuna	3 m
	t-aktub-u	t-aktub-aani	y-aktub-na	3 f
	t-aktub-u	t-aktub-aani	t-aktub-uuna	2 m
	t-aktub-iina	id.	t-aktub-na	2 f

The analysis presented in section 1.8 takes the *y-* as the default prefix and posits two homophonous *t-* prefixes: '2' and 'f.' The reason that *t-* does not appear in the 3 f pl is because 'f' is linked to the suffix position by the affix *-na* 'f pl' before the *t-* 'f' rule applies (cf. 50). The relevant rules are given below:

(94)

	a. t-	2
	b. -na	f pl
	c. t-	f
	d. y-	Elsewhere

There are two important facts to note about the above array. First, there are two homophonous *t-* prefixes. Second, *y-* is the elsewhere prefix. The reason for supposing that *y-* is the elsewhere prefix is that its distribution is not quite a natural class of categories. To state the distribution of *y-* positively, a rule must refer to 3 m *or* 3 pl. By treating *y-* as the elsewhere case, one can instead capture this disjunctive environment as a consequence of discontinuous bleeding: *t-* cannot appear in the 3 f pl because 'f' is discharged by rule (94b) by the time (94c) applies to an input AGR with f pl.

This much I have already discussed. However, observe that while the rules in (94) capture the discontinuous bleeding effect and also avoid the '3 m or 3 pl' disjunction, there is still the peculiarity of having two identical *t-* rules. Not only are there two homophonous *t-* affixes, but the rules introducing these affixes must apply at different times in the derivation, as crucially ordered above.

Because of this homophony, we might expect a reanalysis to take place whereby the *t-* prefix becomes the default case. In this section, I argue that precisely this occurs in Hebrew, Ugaritic, Tell Amarna Canaanite, and South Arabian. All other branches of Semitic (as well as a number of Cushitic languages) treat *y-* or its reflex as the elsewhere case and retain two homophonous *t-* prefixes.

In this regard, consider the Hebrew paradigm in (95) (Lambdin 1971:99):

(95) Hebrew Prefix Conjugation *k t b* 'write'

	singular		plural	
3 m	yi-xtōv	יִכְתֹּב	yi-xtəv-û	יִכְתְּבוּ
3 f	ti-xtōv	תִּכְתֹּב	ti-xtōv-nā	תִּכְתֹּבְנָה
2 m	ti-xtōv	תִּכְתֹּב	ti-xtəv-û	תִּכְתְּבוּ
2 f	ti-xtəv-î	תִּכְתְּבִי	ti-xtōv-nā	תִּכְתֹּבְנָה
1	ʔe-xtōv	אֶכְתֹּב	ni-xtōv	נִכְתֹּב

These forms are much the same as the Classical Arabic forms with one significant exception: the 3 f pl form has the *t-* prefix. To Heb. ti-xtōv-nā תִּכְתֹּבְנָה compare Ar. y .ktub-na.

A similar pattern can be seen in South Arabian. The forms in Mehri can be seen in in section 1.14. The 3 f pl form is tə-rəkz-ən 'they (f) straighten', again with the *t-* prefix. The *t-* prefix occurs in the 3 f pl in Soqoṯri as well (Bittner 1918) and also in Šxawri, as can be seen in the paradigm below from Thomas (1937:28):

(96) Šxawri Prefix Conjugation ayād 'to walk'

	sg	pl
3 m	i-ɣad	i-ɣad
3 f	ta-ɣad	ta-ɣad-an
2 m	t-ɣad	ta-ɣod
2 f	ta-ɣ-it	ta-ɣad-an
1	l-ɣad	na-ɣod

In Šxawri, the 3 f pl form has prefix *ta-* and is homophonous with the 2 f pl form as in Hebrew.[20]

[20]The Šxawri forms have a preverbal particle *ha-* which takes the accent. This particle is omitted for clarity of exposition.

Among the S. Semitic languages, only S. Arabian show the extension of *t-* to the 3 f pl , whereas the Ethiopic S. Semitic languages retain the archaic pattern of Arabic. A simple example is Tigre, where, in an independent development, **y-* has *lə-* as reflex and the dual has been lost (data from Raz 1983)[21]:

(97) Tigre Prefix Conjugation *q n ṣ* 'to get up'

	sg	pl
3 m	lə-qannəṣ	lə-qanṣ -o
3 f	tə-qannəṣ	lə-qanṣ -a
2 m	tə-qannəṣ	tə-qanṣ -o
2 f	tə-qanṣ -i	tə-qanṣ -a
1	ʔə-qannəṣ	ʔən-qannəṣ

The 3 f pl form lə-qanṣ-a 'they (f) get up' does not have the *t-* prefix as in S. Arabian.

The Classical Arabic pattern is also duplicated in Classical Ethiopic (Geʕez) (Lambdin 1978), and the modern Saudi dialects which retain gender in the plural (Prochazka 1988). A similar pattern, although with some other developments, is seen in Syriac (Kaye 1976:147), Chaha (Ethiopic, Ford, 1991:271), and a more conservative dialect of Chadian Arabic, described by Zeltner & Tourneux (1986: 73).

We are thus able to identify two separate geographic areas, S. Arabian and Hebrew, where the 3 f pl form has *t-*, while the rest of Semitic has *y-*.

Consider now the analysis of these special forms. For example, what is the content of the affixes in 3 f pl the Hebrew form ti-xtōv-nā ? Suppose that *ti-* is 'f'. The problem is then how to describe the distribution of the suffix -nā. Suppose it is the allomorph of 'pl' appearing in the feminine. The two rules below will derive ti-xtōv-nā:

(98)

ti-	f
-nā	pl (f)

This analysis is undesirable for two reasons. First, and most obviously, it does not work for the 2nd person forms. Consider the 2 f pl form ti-xtōv-nā, which is homophonous with the 3 f pl form. Here, by analogy to Arabic, one might suppose that there are two *ti-* prefixes, one meaning '2,' the other 'f.' The rules needed to derive the 2 f pl form are:

[21]The prefix *lə-* is historically not the reflex of *y- but rather linked to the Akkadian optative *l-*. One can speculate that *y- weakened and the subordinating particle *l-* , used with this tense/aspect, became attached to the y- initial forms at an early date, not unlike the replacement of *y-* by *b-* in a number of Maghreb dialects of Modern Arabic, owing to the subordinating particle *bi-*.

(99)	ti-	2
	-nā	f pl

For the 2nd person form, the suffix -nā must be a primary exponent of 'f,' whereas for the 3rd person form, it is merely a secondary exponent, i.e. an allomorph of plural appearing when 'f' has been previously discharged. This wrongly implies that there are in fact two homophonous -nā suffixes.

Secondly, observe that in (98) the rules will have to be ordered such that 'f' is realized as *ti-* before the 'pl' suffix is realized. This violates the hierarchy I proposed earlier, whereby gender features are lower than number features and therefore are realized by later rules (other things being equal).

To evade all of these complications, I propose that in Hebrew and in S. Arabian, a reanalysis took place and *t-* became the default prefix of the prefix conjugation. Because there can be at most one elsewhere realization for any given position of exponence, an automatic consequence of the reanalysis is that the former elsewhere prefix *y-* came to be positively identified in some way. We must assume that *y-* was reanalyzed as '3 m' only and not as '3 m *or* 3 pl' as would be necessary to preserve the proto-system distribution.

To see how this works, consider the relevant rules for Hebrew/South Arabian on the left in (100), with the relevant corresponding rules for Arabic/Tigre on the right in (101).

(100)	Heb.	Meħ.	
a.	-î	-i	2 f sg
b.	ni-	nə-	1 pl
c.	yi-	yə-	3 m
d.	?e-	ə-	1
e.	-nā	-ən	pl f
f.	-û	-əm	pl
g.	t-	tə-	Elsewhere
h.	-Ø	-Ø	Elsewhere

(101)	Ar.	Tig.	
a.	n-	?ən-	1 pl
b.	?-	?ə-	1
c.	t-	tə-	2
e.	-na	-a	pl f
f.	-uuna	-o	pl
g.	-iina	-i	f (2)
h.	t-	tə-	f
i.	y-	lə-	Elsewhere
j.	-u	-Ø	Elsewhere

Observe that the two homophonous *t-* prefixes of Arabic/Tigre (101c,h) are cognate with the Elsewhere prefix of Hebrew/Meħri (100g). The elsewhere prefix of Arabic/Tigre (101i) is cognate with the '3 m' prefix of Hebrew/Meħri (100c). These prefixes are reordered automatically by the principles of rule-ordering outlined in the Spell-Out Ordering Hypothesis.

One consequence of the merger of *t-* '2' and *t-* 'f' as the elsewhere prefix is that the suffix rule (101g) must also be reanalyzed. The earlier rule has as its structural description 'f (2),' where (2) needs to be previously discharged for the rule to apply. Since the *t-* rule ceases to discharge '2,' then 'f (2)' cannot apply. To retain its former distribution, the suffix 'f (2)' suffix is reanalyzed as the primary exponent of 2 f sg as in (100a).

The scenario I am proposing may be summarized as follows. The anterior system had *y-* as the elsewhere prefix and two homophonous *t-* prefixes. This homophony made the reanalysis of *t-* as the default prefix a likely occurrence. (I hesitate to say that it *motivated* the reanalysis. One might instead say it *potentiated* the reanalysis). Two branches of Semitic -- Hebrew and S. Arabian -- underwent this reanalysis. An automatic consequence was the extension of *t-* to the 3 f pl form when *y-* came to be positively identified as specifically a *masculine* 3rd person prefix.

It would be natural to expect that such a reanalysis was not immediate nor entire within the relevant speech communities. In earliest Hebrew, for example, certain 3 f pl forms do indeed preserve the anterior state with *y-* as noted by Brockelmann:

> ... there are found in Hebrew sporadic remnants of the formation in i̯ , such as u̯a[-] i̯i̯ēḥamnā 'und sie entbrannten' Gen. 30: 38, u̯a[-] i̯i̯iššarna 'and they [kine] went straight out' 1 Sam. 6:12, i̯aʿa moðnā 'they [four kingdoms] stand [up]' Dan. 8:22 [trans. mine].

Here the cognate of Arabic **y-* can be seen in the early Hebrew prefix i̯-.

The closest relative of Hebrew is Phoenician: these two form the Canaanite branch of Northwest Semitic. Phoenician provides no evidence in this regard since what little evidence remains suggests that feminine gender was impoverished in the plural. In other words, Phoenician, like Egyptian Arabic, seems to have acquired the filter *[pl f], but at quite an early date. For example, Van den Branden (1969:46-7) cites the form *hmt*, normally the 3 m pl pronoun, in a phrase coreferent with a the feminine plural noun šnt 'years.' The common gender 3 pl prefix is *y-*.

As A. Murtonen observes (1967:43), we cannot discern when this development took place in South Arabian either, since the relevant form is not preserved in inscriptions of ancient S. Arabian languages:

> It is a pity indeed that no instances of this person have been preserved in ancient inscriptions [of South Arabian]; on the basis of the present material we must be content to establish that the use of *t-* even in the plural of the 3rd person feminine in all the known S. Arabic dialects is a remarkable distinctive characteristic between S. Arabic and Ethiopic languages...

Ugaritic and the language of the Tell Amarna letters, however, provide decisive evidence in favor of the reanalysis of *t-* as the elsewhere case. Ugaritic is the language of the texts discovered at Rās Šamra, dated to the 14th and 13th centuries B.C.E. The language of the Tell Amarna letters, dating from the 14th century B.C.E. is often called Canaanite, since it displays characteristics of the Canaanite branch -- Hebrew and Phoenician. In all of these languages -- Ugaritic, Hebrew, and Tell Amarna Canaanite, there appears a variability in the prefix for the 3rd person plural.

While all three languages display a more or less regular extension of *t-* to the 3 f pl, each also shows, to varying degrees, the *t-* appearing in the 3 *masculine* plural as well. This further extension has no natural explanation unless *t-* has been reanalyzed as an elsewhere prefix.

In Hebrew, this further extension of *t-* is limited to rather a few examples. These may be treated as the exception, rather than the rule. For example, Gordon (1947b:10) cites the following from Deut. 33:3:

(102) וְהֵם תֻּכּוּ לְרַגְלֶךָ

w-hm t-kw l-rgl-k
and-they T-are.smitten at-foot-thy
'and they are smitten at thy foot'

In the Ugaritic texts from Rās Šamra there is considerably inconstancy in the realization of the 3 m pl. Herdner (1938) discusses the cases of *t-* appearing with masculine subjects in some detail. (Additional discussion appears in Gordon 1947a:63ff.) Among the attested cases, many can be explained away by certain assumptions. For example, the gender of the subject noun may actually have been feminine (103).

(103) t-bl-k ǵr-m m?id ksp
bring.IMPF-you mountain-PL much silver

'the mountains (will) bring you much silver'

Here it might be assumed that ǵr-m 'mountains,' is actually a feminine noun, although its plural ending is typical of masculines. Similarly, one might suppose that in Ugaritic, *inanimate* masculine plural nouns agree as feminines.

Brockelmann hints at such an explanation (1961:§260C, f. Anm.), surmising that the 3 m pl form might arise from the 3 f sg form with the masculine plural suffix. In support of such an idea, he alludes the process of "deflected" agreement in Classical Arabic, whereby nonhuman plural nouns formed by prosodic changes ("broken" plurals) agree as if they were feminine singulars (Thackston 1984:18ff.):

(104) a. bayt-un kabiir-un
house.MASC-INDEF big.MASC.SG.-INDEF.
'a big house.'

b. buyūt-un kabiir-at-un
houses.MASC- INDEF big.SG-FEM-INDEF

'big houses'

Herdner shows that this last hypothesis is untenable, since a good number of examples are attested in which *t-* appears for animals (105) or, more compellingly, specifically human (or divine) masculine plural subjects, (106) below.

(105) ʕl bt ?ab-h nšr-m t-rḫp-n
over house father-his eagles-PL soar.IMPF-PL

"Over his father's house, eagles soar."

In the following examples, the *t-* prefix appears on the verb and a coreferent masculine plural pronominal clitic *-hm* 'their' also appears in the phrase. These examples are particularly compelling insofar as syntactic (rather than interpretive) evidence leads to the conclusion that the subject is masculine.

(106) t-ʕl-n l-mrkb-t-hm
they-climb on-chariot-PL-their

'they climb onto their chariots'

(107) t-š?u ʕil-m r?aš-t-hm
they-raise god-PL head-PL-their

'The gods raise their heads'

A number of examples of 3 masculine *dual* also appear to show the *t-* prefix as well. Consider:

(108) t-š?a-n gh-m w-t-<u>sh</u>-n
raise-DUAL? voice-PL and-cry-DUAL

'They (dual) raise their voices and cry'

In this example, Herdner (1938:81) points out that the subject refers to specifically two messengers who are mentioned by name (Gpn and ?Ugr) earlier in the text.

On the basis of these examples, however subject to possible alternative interpretation, it appears clear that the extension of *t-* to the 3 masculine plural, and quite probably the 3rd dual, was permissible, at least for some stages of Ugaritic. The full paradigm of forms (with vowelled reconstructions) appears below (Gordon 1947a:121):

(109) Ugaritic Prefix Conjugation *m l k* 'to rule'

singular	dual	plural	
y-mlk	y/t -mlk-n	y/t -mlk-n	3 m
t-mlk	t-mlk-n	y/t -mlk-n	3 f
t-mlk	t-mlk-n	t-mlk-n	2 m
t-mlk-n	id.	t-mlk-n	2 f
?-mlk	unattested	n-mlk	1
*y-amluk-u	*y/t -amluk-aani	*y/t -amluk-uuna	3 m
*t-amluk-u	*t-amluk-aani	*y/t -amluk-na	3 f
*t-amluk-u	*t-amluk-aani	*t-amluk-uuna	2 m
*t-amluk-iina	id.	*t-amluk-na	2 f
*?-amluk-u	??	*n-amluk-u	1

The extension of *t-* appears much more frequently in Tell Amarna Canaanite. Herdner (1938:76) cites:

(110) a. t-idū 'they know' (105,36)
b. t-idūkū 'they killed' (75,33)
c. t-ilkū 'they take' (84,32)
d. t-adinūni 'they gave' (126, 63)

Brockelmann (1961: §260C, f. Anm.) cites in addition:

(111) a. t-ašpurūna 'they send' (B. 58,123)
b. t-uballiṯūna 'they (G. man) supply provisions' (L.13,56)
c. lā t-ugammarūna 'so that they do not destroy , G. damit sie nicht vernichten' (L. 49, 25).

The frequency of such usage suggests that for Tell Amarna Canaanite, the extension of *t-* approached the rule rather than the exception.

According to my hypothesis, the Canaanite languages underwent a reanalysis whereby *t-* became the elsewhere prefix. Only by assuming this can it be explained how *t-*, formerly conditioned by '2' or 'f,' came to appear in a category that was neither 2nd person nor feminine, and indeed in a set of forms (all 2nd and 3rd person forms except 3 m sg), which form no natural class.

One can explain the variability in the appearance of *t-* by assuming that *y-* was positively identified sometimes as '3 m sg' rather than '3 m.' In other words, the following rules applied in cases where *t-* appears in the 3 m dual and 3 m pl:

(112) a. y- 3 m sg
b. n- 1 pl
c. ʔ- 1
b. t- Elsewhere

Reanalysis of *y-* as specifically singular allows *t-* as the Elsewhere prefix to appear automatically in the 3 m pl and 3 m dual in Ugaritic and Tell Amarna Canaanite. This rule system differs only minimally from the system in which *t-* does not extend to the masculine forms (Hebrew, some Ugaritic), where *y-* is '3 m' only (cf. rule system (100)).

In this way, the extension of *t-* receives a simple explanation, and a long-standing puzzle of Semitic historical morphology is solved. One need no longer appeal to "false analogical formation," (cf. Brockelmann 1961 §260C, f, Anm.) or other pseudo-explanations.

From a theoretical point of view, the analysis I have presented crucially relies on the notion that some affix may be an Elsewhere affix, devoid of feature content, and realized only if no other more specific affix may fill a position of exponence. Moreover, the choice of which affix is the Elsewhere affix is subject to historical change.

Lexical affix-based models require that each affix have some feature content as part of its lexical representation, and, as I have already discussed, cannot permit the notion of an Elsewhere affix. Therefore, the lexical affix-based model will be unable to express the extension of *t-* in Ugaritic and Tell Amarna Canaanite as a natural historical process. For this reason, I assume that the lexical affix model is explanatorily inadequate without further modification.

1.16 The Loss and Reappearance of *t-* 'f' in Akkadian

Within Semitic, Old Akkadian (2500-2000 B.C.E.) displays the archaic Semitic pattern of prefixes. After 2000 B.C.E. Akkadian split into two dialects: Babylonian, spoken in the southern region of Mesopotamia, and Assyrian, spoken in the northern region. Babylonian shows an interesting development: instead of showing the regular *t-* in the 3 f sg, Babylonian documents regularly have *i-* in both genders in the 3rd person (von Soden 1969:69).[22] An example paradigm from Babylonian is shown in (113), (Reiner 1966: 70).

(113) Babylonian Prefix Conjugation (preterite) *p r s*

	sg	plural
3 m	i-prus	i-prus-u
3 f	i-prus	i-prus-a
2 m	ta-prus	ta-prus-a
2 f	ta-prus-i	ta-prus-a
1	a-prus	ni-prus

For Old Akkadian 3 f sg ta-prus, Babylonian regularly has i-prus. Assyrian, however, continues the Old Akkadian pattern. The loss of *ta-* 'f' was of course not immediate, and there are some instances of archaism in Old Babylonian (von Soden 1969:99).

[22]Some verbal stems show preradical *u-* in place of *i-*. The discussion here will be limited to one class of verbal stems.

Recall that according to the analysis I have presented, the oldest Semitic system contains two homophonous prefixes *t-* '2' and *t-* 'f.' I argued in the last section that this homophony made the reanalysis of *t-* as the elsewhere prefix a plausible development. Inversely, the development seen in Babylonian suggests that *t-* 'f' was lost and that **y-* (Akk. *i-*) spread to fill the place of *t-*.

Simple loss of *ta-* 'f' suffices to derive the change from Old Akkadian to Babylonian. However, we should like to discover some way of expressing this loss as a natural development. Consider first the affixes which derive the Old Akkadian forms:

(114) Old Akkadian (and Assyrian)

a. ni- 1 pl
b. a- 1
c. ta- 2
d. ta- f
e. i- Elsewhere

The Babylonian forms can be derived by the same rules except (114d).

However, a more explanatory scenario can be constructed to express the change. Suppose *ta-* was reanalyzed as the Elsewhere prefix, causing *i-* to be positively identified. I argued in the last section that precisely this reanalysis took place in the Canaanite branch. If *i-* were positively identified as '3' in Babylonian instead of as '3 m' as in Hebrew, or '3 m sg' as in Ugaritic and Tell Amarna Canaanite, then *i-* would spread automatically to the 3 f sg. To see this, consider the rules that would result in such a case:

(115) Standard Babylonian

a. ni- 1 pl
b. a- 1
c. i- 3
d. t- Elsewhere

For a 3 f sg input, (115c) will apply deriving i-prus in (Standard) Babylonian. Now the elsewhere rule will *de facto* apply only to the 2nd person forms.

In a further development, von Soden (1969:99) notes that in the last phase of New Babylonian, referred to as Late Babylonian (Spätbabylonisch), attested from 600 B.C.E., *ta-* reappears in the 3 f sg. In other words, the Late Babylonian system is identical in this respect to the Old Akkadian and Assyrian systems. Because late Babylonian came heavily under the influence of (Classical) Aramaic, von Soden attributes this development to Aramaic influence. Aramaic retains the archaic Semitic prefix pattern.

To understand this final development, consider first the Babylonian rule system in (115). I will assume, following C. Watkins (1962), that a category marked 3rd person singular, or as here, 3rd person alone, tends automatically to be reanalyzed as the elsewhere or least marked case in any morphological system. If this is so, one should expect that the Babylonian system should be relearned by successive generations as:

(116) New Babylonian

a. ni- 1 pl
b. a- 1
c. t- 2
d. i- Elsewhere

Note that the *t-* prefix is now only '2.' During the intermediate stage of Standard Babylonian, *t-* no longer appeared in the 3 f sg, and so, when positively identified, it was reanalyzed only as '2.'

It appears that under the influence of Aramaic, New Babylonian (re-)acquired the archaic *ta-* 'f' prefix. Adding this rule to (116) automatically derives the Late Babylonian system, which is in fact identical to the Old Akkadian and Assyrian system in (114).

To summarize, Old Akkadian shows the archaic prefix pattern of Semitic. As such, it displayed the two homophonous *t-* prefixes. The loss of *t-* in (Standard) Babylonian may be seen as the reanalysis of *t-* as the Elsewhere prefix followed by reanalysis of **y-* (Akk. *i-*) as '3' only. We thus discern three responses to the reanalysis of *t-* as the elsewhere prefix:

(117)

Proto-Semitic:	*t-* '2'; *t-* 'f' ; *y-* Elsewhere
Babylonian:	*i-* '3' ; *t-* Elsewhere
Hebrew:	*y-* '3 m' ; *t-* Elsewhere
Ugaritic / Canaanite:	*y-* '3 m sg' ; *t-* Elsewhere

In each case, upon losing its status as Elsewhere prefix, **y-* came to be positively identified in a different way. To retain the archaic distribution, *y-* would need to be positively identified disjunctively: '3 m *or* 3 pl.' Babylonian simplified this disjunction to '3,' and Hebrew to '3 m.' Ugaritic / Canaanite went further and reanalyzed it as '3 m sg.' The more specifically *y-* was positively identified, the more the elsewhere prefix *t-* spread through the paradigm.

Finally, I argued that, owing to the universal tendency for '3' or '3 sg' affixes to be reanalyzed as elsewhere affixes (deriving from Watkins' observation), the Babylonian system changed such that *t-* was positively identified as '2' and *i-* again became the elsewhere prefix. The reappearance of archaic *t-* 'f,' in Late Babylonian under Aramaic influence can then be analyzed simply as the gain of the *t-* 'f' affix within this system.

The hypothesized developments are shown below:

(118)

Old Akkadian >		Babylonian >		New Babylonian >		Late Babylonian	
ta-	2	i-	3	ta-	2	ta-	2
ta-	f					ta-	f
i-	Els.	ta-	Els.	i-	Els.	i-	Els.

1.17 Feature-changing Syncretism

The analysis presented up to this point has assumed that the two *t-* prefixes seen in Classical Arabic or Old Akkadian are separate but homophonous affixes. In this section I detail another possible analysis in which this relationship is treated not as an accidental homophony, but rather as a deeper syncretism.

What is required is a *feature-changing* rule operating on morphosyntactic representations before the phonological realization of affixes. Consider the following such rule:

(119) 3 --> 2 in env. f {sg, dual}

This rule changes a 3rd person feature to a 2nd person feature in the environment of f and singular or dual.[23] This rule precedes the rules in (120):

(120)
a. n- 1 pl
b. ʔ- 1
c. y- 3
d. t- Elsewhere

The effect of rule (119) is essentially to bleed the application of (120c) from the 3 f sg and 3 f dual forms. For example,an input AGR 3 f sg will be changed to 2 f sg by (119). When the rules in (120) apply, rule (120c) will not apply and the form will receive the elsewhere prefix *t-* by rule (120d).

[23]In a more complete analysis, this rule would be insufficient. Note that if 3 --> 2 in the f sg, then we should expect 3 f sg to also show the f (2) suffix *-iina*, contrary to fact. To remedy this, rule (119) must change 3 f to 2 m. One consequence of this is that the value 'masculine' must be introduces into the morphosyntactic representations. This value otherwise appears to be absent from the Arabic system as a marked value, and needs to be invoked here only to avoid *-iina* after (119) applies.

By assuming the feature changing rule in (119), it is no longer necessary to assume that Proto-Semitic (and hence Classical Arabic) system had two homophonous *t-* affixes. On the one hand, such an analysis is attractive because it treats the relationship between the feminine nonplural and 2nd person as a deep syncretism rather than a phonological accident (homophony). Because *t-* appears in both the 3 f sg and the 2 categories in Cushitic as well as Semitic, it appears that this relationship is indeed quite archaic.

Furthermore, there seems to be an independent need for feature-changing syncretism rules in morphology. For example, in Russian the genitive of a masculine or plural noun is either the nominative if the noun is inanimate, or genitive if the noun is animate.[24] Examples of masculine singular nouns are given in (121).

(121) Russian masculine noun syncretism

	nom.sg.	gen. sg.	acc.sg.
anim. 'writer'	*pisátel'* писáтель	*pisátel'a* писáтеля	*pisátel'* писáтеля
inan. 'museum'	*muzej* музей	*muzeja* музея	*muzej* музей

Halle (1991) has proposed the rules in (122) to derive these syncretisms directly:

(122) a. ACC --> GEN / masc. or plural animate
b. ACC --> NOM / masc. or plural inanimate

If Halle is correct in supposing that feature-changing rules like these are permitted in morphology, then there is no *a priori* reason to suppose that similar rules are not responsible for the *t-* syncretism in Semitic.

On the other hand, other factors suggest that the *t-* syncretism is in fact a homophony. First of all, it should be observed that the rule as formulated in (119) is a somewhat specific rule. In particular it is constrained not to apply in the 3 f pl. Recall that on the discontinuous bleeding analysis, 3 f pl is special because 'f pl' has its own affix, *-na*, which bleeds 'f' from the prefix position, allowing elsewhere *y-* to appear. The feature-changing analysis cannot capture this fact except by stipulation -- albeit a negative sort of stipulation, that is, the failure of (119) to apply in the 3 f pl. In the discontinuous

[24]For a survey of this syncretism in Slavic and its historical development, see Huntley (1980). For a thorough discussion of this syncretism, particularly in numeral phrases, see Babby (1987).

bleeding/ homophony analysis the failure of *t-* to appear in the 3 f pl follows automatically and without stipulation.

Second, it is conceptually bizarre that rule (119) changes the 3 category to a category -- 2 -- which has no explicit spell-out. The effect of (119) is to bleed forms into the elsewhere class. In general, where feature-changing rules are well motivated, bleeding is into an explicit category with a distinct realization.

Third, one must consider how a rule like (119) could have arisen historically. Because *t-* has uses both as a 2nd person marker and as a feminine marker (nominal *-at*, for example), the standard assumption has been that the two *t-* prefixes were originally merely homophonous. Therefore, the homophony analysis must be correct at some stage. The final conclusion may be that both the homophony and the feature-changing analyses have merit and are applicable to different stages of Afroasiatic. At the deepest level, both analyses derive the facts by a bleeding relation, whether by feature-changing or by discharge at different string positions.

The historical changes detailed in the last two sections can receive accounts in the feature-changing analysis. For example, the loss of *t-* 'f' in Babylonian can be understood as loss of rule (119). The reappearance of *t-* 'f' in Late Babylonian is the addition of rule (119), borrowed from Aramaic. The spread of *t-* in Hebrew is the loss of 'sg/dual' form rule (119). For now, I leave open the question of to what degree feature-changing analyses are to be permitted. In Chapter 3, I will show that the conversion of duals in trials to plurals in Nunggubuyu is not a structure-changing syncretism but rather a two-step process of deletion and insertion, which is another possible avenue of analysis for what appear to be structure-changing processes.

Because feature-changing rules are so powerful (any category could become any other in principle), they should either be impossible or highly costly within an evaluation of grammatical complexity. Lumsden (1992: 472-3) has in particular argued against feature-changing syncretism rules on general grounds:

> If these rules change the feature-values of the underlying positions, then the underlying distribution of features has very little relation to the semantic/syntactic environment. That is, there is no longer any systematic relationship between the distribution of features in syntactic positions and the semantic/syntactic significance that is associated with feature labels. Ultimately, this complex of rules has no motivation other than to mechanically describe the distribution of forms in the surface structure.

A number of counterarguments might be presented to Lumsden's complaint. First, after syntax and within a component of Morphology, there is no reason to suppose that morphosyntactic labels still have any semantic or syntactic 'significance.' In other words, if a feature ACC is changed to GEN in Morphology, then the label GEN at that position has only morphological significance *at that point in the derivation.* Second, there is in fact a 'systematic' relationship of features (representing semantic/syntactic environments) to their changed-feature counterparts (representing morphological signals), and that relationship is the set of feature-changing rules themselves. To quarrel with this relationship is to deny that there can be feature-changing rules, and that is precisely the question at hand. Finally, we must carefully examine what constitutes the 'motivation' of a set of rules. Here again, the question does not hinge on descriptive adequacy, since feature-changing analyses are no more descriptively adequate that others: the question is rather what properties of a feature-changing analysis would make it a mere 'mechanical description.'

Clearly, arguments for or against feature-changing syncretism must address the question of learnability, as theoretically embodied in an evaluation of cost. A feature-changing syncretism analysis is indeed very abstract and complex, insofar as the systematic relationship between semantic and syntactic environments and morphological signals is *non-transparent* and mediated by feature-changing rules. The questions becomes whether such abstraction and complexity are admissible, by which it can only be meant, learnable. Feature-changing analyses in morphology, like those in phonology, must be committed to the idea that features are mental entities figuring in the computation of surface forms and not somehow mere transparent reflexes of surface forms.

In light of the data from Russian, it may be too hasty to conclude that feature changing syncretisms are unlearnable. A more cautious conclusion is that they are learnable only given a sufficient stimulus, that is to say, such rules are highly costly. If alternative analyses exist, they are presumably less costly and therefore more likely to reflect speaker's knowledge of morphology. On these grounds, I will not advocate the feature-changing analysis for the Semitic forms, since I have presented what I believe to be a less costly homophony analysis.

1.18 Discontinuous Bleeding in the Berber Prefix Conjugation

More evidence for the discontinuous bleeding relation comes from the prefix conjugation in Berber. The Berber languages form one branch of Afroasiatic and show a separate development in their prefix conjugation.

Berber is still spoken in a number of non-contiguous enclaves in Morocco and Algeria. Tuareg, spoken in the Saharan region from Mali eastward to Sudan, is related to Berber and is considered part of the Berber dialect chain.

Berber differs from Arabic in that its verbs have no autonomous morphological structure: in other words, no position of exponence is required by independent conditions on well-formedness. Such systems will be called Free Licensing systems, since affixes may appear without being licensed by positions in an autonomous template. I will have much more to say about Free Licensing in chapter 3.

1.18.1 The Prefix Conjugation in Tamazight

The prefix conjugation for the Tamazight dialect of Central and Southeastern Morocco is shown below (Abdel-Massih 1971):

(123) Tamazight Berber Prefix Conjugation *dawa* 'cure'

	sg	pl
3m	i-dawa	dawa-n
3f	t-dawa	dawa-n-t
2m	t-dawa-d	t-dawa-m
2 f	id.	t-dawa-n-t
1	dawa-ɣ	n-dawa

Certain traits of the Berber conjugation are analogous to the Semitic one, whereas others are quite different. For example, the 1 pl prefix is still *n-*, and *t-* marks both 3 f and 2 again. Other affixes are different: whereas Semitic has the same affix in the 3 m pl and 2 m pl (Ar. *-uuna*, Heb. -û), Berber has *-n* in the 3 m pl and *-m* in the 2 m pl. The suffixes in the singular, namely *-d* in the 2 m sg and -ɣ in 1 sg, are without parallel in the Semitic prefix conjugation. Each of these discrepancies can be explained by assuming that the Berber prefix conjugation is actually a mixed one, showing some characteristics of the Semitic prefix conjugation, and some of the Semitic suffix conjugation (Prasse 1963). The 2 sg form t-dawa-d shows both the archaic **t-* prefix and a reflex of the suffix conjugation *-*tV*. The suffix -ɣ is a reflex of *-kū with cognates in the Semitic suffix conjugation (Ar. -tu, Meħ. -k).

A general fact about Berber verbs is that gender is distinguished in the 2nd person only in the plural. In other words, gender is impoverished in the 2 sg categories. Observe that the 2 m sg and 2 f sg forms are identical in (123).

To express this property, I propose that Berber verbs have the following filter:

(124) *[2 sg f]

Because of the feature hierarchy, the filter in (124) conditions the deletion (delinking) of 'f' from any input AGR '2 sg f.' It is significant that this filter can be expressed only if 'sg' is visible at Morphology, that is, if 'sg' is not underspecified at this level. It is predicted then that 'sg' should in fact be visible in the rules/affixes of the system. We will see that this is the case, and crucially so.

Another fact of interest about the Berber forms is that here the input AGR splits into three positions of exponence instead of two as in Semitic. For example, consider the 2 f pl form t-dawa-n-t. This form can be analyzed as

(125) t-dawa-n-t 'you (f, pl) cured'
2-cure-pl-f

It is important to observe that, as in Semitic, both number and gender (as well as person) may appear in the prefix position, as in

(126)
a. n-dawa 'we cured'
1 pl-cure

b. t-dawa 'she cured'
f sg-cure

Recall that in Semitic, person features never appear in the suffixal position. The same cannot be said of Berber:

(127)
a. dawa-ɣ 'I cured'
cure- 1

b. t-dawa-m 'you (m, pl) cured'
2-cure- pl m (2)

All three types of AGR features may appear in the suffix position. Berber shows an even more dramatic case of AGR splitting.

In addition to a 3-way splitting of AGR features, Berber differs from Arabic in regard to whether the positions occupied by the affixes are obligatory. In Arabic, we saw that a prefix and a suffix were obligatory affixes in the prefix-conjugation (except the imperative). In Berber, none of the three positions -- prefix, first suffix, and second suffix -- is obligatory: each may be 'occupied' by a Ø:

(128) a. i-dawa-Ø-Ø
m sg-cure-Ø-Ø 'he cured'

b. Ø-dawa-ɣ-Ø
Ø-cure-1-Ø 'I cured'

The above examples show that each of the three positions may be empty.

The crucial difference between the Arabic -Ø suffix (jussive) and the Berber Øs above is that the Arabic Ø occupies an obligatory position and is contentful. By this it is meant that there is an affix Ø in Arabic which discharges (spells out) the feature [jussive] and fulfills the autonomous morphological requirements of the prefix conjugation. The Ø "affixes" in Berber are not contentful and do not discharge any features, as we will see. Associated with this is the fact that the three positions in Berber are not obligatory: if no positively-identified affix fills the prefix position, then the prefix position need not be discharged (filled with) some Elsewhere affix as in Arabic. The zeroes appearing in (128) are thus really nulls with no linguistic content or reality.

Systems such as Berber differ from those in Arabic by a parameter which I identify as [± Autonomous Licensing]. When affixes must be licensed by autonomous well-formedness conditions, as in Arabic, then the M^0 with which they are associated is [+Autonomous Licensing]. If the M^0 splits indefinitely into a number of *non-obligatory* positions-of-exponence, as in Berber, then this M^0 is [-Autonomous Licensing]. I will have more to say about this in Chapter 3.

As we will see, it is no accident that where there are two suffixes, the first one realizes 'pl' and the second 'f.' This follows because the rule introducing *-n* 'pl' precedes *-t* 'f' by the hierarchy of features. In Free Licensing systems, the unmarked situation is for the order of attachment of affixes to be the same as the hierarchy of affixes when they are competing for the same position-of-exponence. I return to this matter again in chapter 3.

1.18.2 An Analysis of Tamazight

The following set of rules derives the Tamazight pattern:

(129) Tamazight Prefix Conjugation

a.	n-	1 pl
b.	-ɣ	1
c.	t-	2
d.	-m	pl m (2)
e.	i̯-	sg m
f.	t-	sg f
g.	-d	sg (2)
h.	-n	pl
i.	-t	f

discontinuous bleeding

In the rule/affix system above, there are four cases of discontinuous bleeding; the lines above show the relevant rule orderings. Each of these will be discussed in turn.

The first case is shown by the ordering (129a) < (129b). The suffix -ɣ '1' need not be specified as specifically 'sg,' since any nonsingular 1st person form will discharge '1' by the previous rule (129a).

The second case of discontinuous bleeding is (129a) < (129h). This is a relationship similar to the bleeding of Ar. *-uuna* 'pl' by *n-* '1 pl.' In Tamazight, the plural suffix *-n* is bled by the prefix *n-* '1 pl.' Note that because *-n* 'pl' appears in the unnatural class 2 f pl and 3 pl, it is best to analyze *-n* as merely 'pl.'

The third case is (129d) < (129h). This case is less obvious. Because Berber permits more than one AGR suffix, as evidenced by (125) t-dawa-n-t 'you (f pl) cured,' unless discharge of 'pl' is invoked, there is no reason why both (129d) and (129h) could not both apply to a form. The ill-formed verb

(130) *t-dawa-m-n
2-cure-m pl (2)-pl

would result. As things stand, (130) could never arise since 'pl' will be discharged by (129d) before (129h) applies.

The last case in (129f) < (129i): *t-* ‘f sg’ before *-t* ‘f.’ Consider the 3 f sg form t-dawa. If ‘f’ is not discharged at the prefix position, rule (129i) should supply *-t* to the (second) suffix position, yielding *t-dawa-t..

On the whole, Tamazight provides four cases of discontinuous bleeding and an even more radical splitting of AGR than is seen in Semitic. Even so, this conjugation can be derived without special stipulations by the 9 rules in (129).

It is important to note that because there are no obligatory positions in Berber, each spell-out rule does not discharge (fill) a position as in Arabic. Therefore, there can be no competition for ‘slots’ in autonomous morphological structure, and the affixes introduced by the rules in (129) are never disjunctive with respect to particular positions. However, because the rules continue to discharge *features* of the input AGR, discontinuous bleeding still prevents more than one affix to appear realizing the same property as in (130).

1.18.3 A Word-Paradigm Analysis of Tamazight

It is possible to capture all of the discontinuous bleeding relationships discussed above in the Extended Word Paradigm Model, but at some cost. It will be instructive to examine how such an analysis will work. We consider first an analysis in which the discharge effect is limited to the rule block only; next, this analysis will be compared to one in which rule disjunctions across rule blocks are permitted in a limited way, after the proposal of Anderson (1986). It will be shown that this latter analysis differs in fact only minimally from the current proposal.

Recall that in the earlier Extended Word-Paradigm model, the discharge effect is limited to the rule block only. Therefore, in constructing an analysis of the Berber facts, each discontinuous bleeding relation must be a rule disjunction within a rule block. For example, because *n-* ‘1 pl’ and *-n* ‘pl’ are disjoint, they must occur in the same block of rules.

Inversely, where I have posited allomorphy relations, i.e. features in parentheses, EWP must put the relevant rules in separate blocks. For example, because *t-* ‘2’ and *-d* ‘sg (2)’ are not disjoint, that is, they co-occur for example in the form *t-dawa-d* ‘you (sg) cured’, then these rules must appear in separate blocks.

These two considerations restrict rather significantly the EWP analogue of (129). For simplicity, I will refer to the rules by the letter of (129) in which they occur. The following are disjunctive by discontinuous bleeding and must therefore be in the same block: {a,b}, {a,h}, {f,i} and {d,h}. Because {i} cooccurs with {h} in the 2 f sg form, {i} and {h} must be in separate blocks, and therefore {f,i} must be in a separate block from the other disjunctive pairs. Because {c} cooccurs with {d}, {g}, {h}, and/or {i} in the various 2 forms, {c} must be separate from {d}, {h}, and {i}. The only possible solutions are {{c}, {a,b,d,h}, {e,f,g,i}} and {{c,e}, {a,b,d,h}, {f,g,i}}. These groupings differ minimally by which rule block {e} appears in. Let us consider only the former grouping, displayed below:

(131) Tamazight Prefix Conjugation (early EWP Analysis)

Block I:		X {2} --> t+X	(c)
Block II:	a.	X {2 m pl} --> X+m	(d)
	b.	X {1 pl} --> n+X	(a)
	c.	X {1} --> X+γ	(b)
	d.	X {pl} --> X+n	(h)
Block III:	a.	X {2 sg} --> X+d	(g)
	b.	X {m sg} --> i+X	(e)
	c.	X {f sg} --> t+X	(f)
	d.	X {f} --> X+t	(i)

I showed in section 1.3 that EWP could not capture the discontinuous bleeding effects in Arabic without encountering a paradox. These paradoxes were based on the distribution of elsewhere affixes in the Arabic conjugation. Because the elsewhere affix in Berber is uniformly -Ø -- that is to say, there are autonomously required positions-of-exponence, the same arguments cannot be constructed for Berber.

However, a comparison of the two systems (129) and (131) is revealing. The discontinuous bleeding analysis requires only one block of rules whereas EWP requires three blocks. From the point of view of learning the forms of the system, one must assume on the EWP analysis that one must learn both the rule *and* the block it occurs in. For example, it does not suffice to know only that Berber has a rule prefixing *t-* to 2nd person forms, one must also know that this rule occurs in Block I and therefore is not disjunctive with the other rules of the system. The Block which a rule belongs to is not in any way correlated with its position of exponence either: both prefixes and suffixes occur in both blocks II and III, so the property of belonging to a certain rule block is indeed highly abstract.

In contrast, the analysis I have given in (129) requires only that each affix be learned associated with its feature content. The rules introducing the affixes apply in one block for each M^0 (here, AGR), with all rules being ordered by the feature hierarchy. Crucially, however, I assume that an affix is the primary exponent of some features and the secondary exponent of others. I have notated in parentheses the features for which the affix is a secondary exponent.

Recall now from section 1.3 that on the proposal of Anderson (1986, 1992), disjunctions are permitted across rule blocks if the blocked rule's structural description contains a (nonnull) subset of the features contained in the blocking rule's structural description. This effectively mirrors the effect of feature discharge in the current proposal. Given this modification, the Berber rules can be expressed as follows:

(132) Tamazight Berber Prefix Conjugation (Revised EWP analysis)

Block I:
a. X {1 pl} --> n+X
b. X {2} --> t+X
c. X {m sg} --> i+X
d. X {f sg} --> t+X

Block II:
e. X {2 m pl} --> X+m
f. X {2 sg} --> X+d
g. X {1} --> X+ɣ
h. X {pl} --> X+n

Block III:
i. X {f} --> X+t

Certain disjunctive relations now hold across rule blocks owing to the subset condition: in the following pairs, the first rule blocks the second: {a, g}, {a, h}, and {d,i}. However, recall that the blocking rule must precede the blocked rule; therefore the following pairs will not be disjunctive: {b, e}, {b, f}.

Comparing the revised EWP analysis (132) with the proposal in (129) reveals some striking similarities. Both have essentially the same rules, but (132) has three rule blocks, corresponding to three position classes; the proposal in (129) has only one block of rules. What is accomplished on my analysis through secondary exponents is accomplished in the revised EWP framework by situating rules in separate blocks. For example, consider the -*m* rule/affix. On my analysis, this affix has the content 'pl m (2),' which delays its application until after the *t*- '2' rule has applied. This follows because -*m* is by hypothesis only a secondary exponent of '2.' The highest feature on the hierarchy for which -*m* is a primary exponent is 'pl'; since 'pl' is lower than '2' on the hierarchy of features, the -*m* affix attaches after the *t*- affix. On the EWP analysis, the -*m* rule follows the *t*- rule only by stipulation: the two rules are in separate blocks, and the latter rule's block applies first. It is an interesting consequence that the rule blocks in (132) correspond to position classes, but it is essential to observe that this need not be the case. As a result, the situation of each rule in a particular block is *entirely a matter of stipulation* on the EWP model. On the present proposal, there is only one rule block, corresponding to the M^0; all rule orderings follow by principle within this one block. The effect of delaying a rule's application to a later block in the EWP model is accomplished in my analysis by the machinery of secondary exponence.

1.18.4 Dialect Differences within Berber

The remaining dialects of Berber and most of the Tuareg dialects differ minimally from the Tamazight pattern.

One pattern of dialectal difference is the phonological realization of the suffix -*d* 'sg (2).' According to Basset (1952), this *d* is emphatic in the Taqbaylit / "Grande Kabylie" dialect (Northern Algeria), and is devoiced unemphatic -*t* in the Tachelhiyt ("Tachelhait / Shilha") dialect (S.W. Morocco).

Tamazight is unusual among the dialects in that it has the suffix -*m* only in the masculine of the 2 pl. Tachelhiyt (Applegate 1958) has -*m* in both genders in the 2 pl, and a similar pattern is seen in Tuareg (Prasse 1963, 1973). Compare the plural forms in (133):

(133)

	Tamazight	Tachelhiyt
2 m pl	t- -m	t- -m
2 f pl	t- -n-t	t- -m-t
3 m pl	-n	-n
3 f pl	-n-t	-n-t

The restriction of -*m* to specifically 2 m pl is apparently an innovation of Tamazight. To derive the earlier pattern, rule (129d) need merely be simplified as

(134)	d. -m	pl (2)

The extension of *-n* to 2 f pl in Tamazight might be mere assimilation to the following *-t* 'f.' Or it may be the result of reanalysis of two-piece *-n-t* as one-piece *-nt* 'f pl' in this dialect. If *-n-t* is reanalyzed as a single affix, then the resultant system of plural suffixes is:

(135) Reanalyzed Tamazight plural suffixes

a. -nt	f pl
b. -m	pl (2)
c. -n	pl

Each rule bleeds the next rule(s). One by-product of the reanalysis of *-n-t* as a single affix is that *-t* no longer occupies a second suffix position and AGR need not split into three positions of exponence but only two. This simplification may be seen as offsetting the partial homophony of *-nt* 'f pl' and *-n* 'pl' in (135).

1.18.5 The Extension of -ɣ in Three Tuareg Dialects

Basset (1952) first drew attention to a peculiar replacement of –d 'sg (2)' by –ɣ '1' in the 2nd person singular forms in the Iouellemeden dialect of Tuareg. Prasse (1971) identifies eight dialects of Tuareg as differentiated by seventeen isoglosses. The –ɣ for –d region comprises three of the eight dialects, namely (after Tuareg designations) Tayärt (central Niger), Tämәsgәrәst (south-central Niger), and Täwәllәmmәt tan dәnnәg (i.e. Eastern Täwәllәmmәt, the Iouellemeden dialect spoken in western Niger and part of Mali).

Consider the following pairs (Prasse 1971:204). Prasse's *ä* = /ʌ/ :

(136)	Conservative		Innovating	
	1 sg	äls-äɣ	äls-äɣ	'I get dressed'
	2 sg	t-äls-äd	t-äls-äɣ	'you (sg) get dressed'

Aside from this isogloss, the general pattern of prefix conjugation affixes in Tuareg is isomorphic to that of Berber generally and requires no other comment (Prasse 1973).

To express this extension, we will have to construct a system in which -äɣ '1' extends automatically to the 2nd person singular. At first glance, such an extension seems highly improbable. However, it need not be.

An examination of all the rules in (129) reveals a certain pattern of distribution. The suffix -(ä)ɣ '1' is the only instance where a person AGR feature is linked to a suffix. In the innovating dialects of Tuareg, person features became confined solely to the prefix position and -(ä)ɣ '1' was reanalyzed as 'sg.'

It is important to observe that this reanalysis will have no overt effect of the system in (129). Consider the rule system after -(ä)ɣ becomes just 'sg.' (Certain phonological changes are noted for Tuareg, following Prasse (1973:9)):

(137) Tuareg Prefix Conjugation

a.	n-	1 pl
c.	t-	2
d.	-(ä)m	pl m (2)
e.	i-	sg m
f.	t-	sg f
g.	-äd	sg (2)
h.	-äɣ	sg
i.	-(ä)n	pl
j.	-ät	f

discontinuous
bleeding

All the singular categories besides '1 sg' will have 'sg' discharged by a rule applying before rule (137h). The 2 sg categories will undergo (137g), and the masculine and feminine 3rd person categories will undergo (137e,f) respectively. The 1 sg category is the only one which will have 'sg' undischarged when (137h) applies, and so -äɣ can have 'sg' only as its content.

In this way, it is possible to see that the suffix -äɣ is amenable to two different analyses without affecting the surface forms: either it is '1' or it is just 'sg.' The reanalysis as 'sg' is the first step in the extension of -äɣ in the innovating dialects. These dialects then lost *-äd*, rule (137g). The singular suffix -äɣ automatically appeared in the 2 sg forms, since -äɣ had lost its specifically 1st person content.

Thus we are able to give a remarkably simple account of what otherwise appears to be a peculiar analogical formation. The story I have constructed crucially depends on the notion of discharge. In order for -äɣ to be reanalyzed as 'sg,' all other sg categories must be discharged by earlier rules. In particular, for the 3rd person categories, 'sg' is discharged at the prefix position as either i̯- or *t*-.

To express this same change, the early Word-Paradigm analysis given in (131) would need to say that the rule suffixing -äɣ was reanalyzed as being conditioned by 'sg' rather than by '1,' and *moreover* moved from Block II to Block III so as to become disjunctive with the rules in Block III. Other accounts of the change might be constructed, such as conditioning -äɣ by specifically [-3] categories, or other postulation of extra features. On the revised EWP analysis (132), the effects follow roughly as in my proposal: rule (132g) is reanalyzed as being conditioned by 'sg,' and then rule (132f) is lost with the correct results. On my analysis, the change is a vacuous reanalysis of -äɣ followed by the loss of the affix *-äd* 'sg (2).'

In sum, the Berber facts demonstrate that the notion of feature discharge across position classes (captured in the revised EWP analysis by means of disjunctions across rule blocks) must be recognized as more than a chance occurrence; it has both a synchronic validity and permits the explanation of 'analogical' change. Regardless of the ultimate theoretical analysis, both the Berber facts of this section and similar effects seen in Arabic establish discontinuous bleeding as a robust phenomenon within inflectional systems.

1.19 Summary

In this chapter I drew attention to the phenomenon of discontinuous bleeding in the Afroasiatic prefix conjugation. Where discontinuous bleeding occurs, an affix realizing a feature F will not appear if F has been realized by another affix at some other position of exponence in the form. This rarely discussed phenomenon motivates a break with all existing models of inflectional morphology.

I contrasted three existing models of inflectional morphology, showing how each is inadequate to explain the prefix-conjugation facts without further modification.

On the basis of the discontinuous bleeding facts, I developed a theory whereby features in an input morphosyntactic representation associated with a syntactic position (an M^0) are discharged by rules which rewrite positions in autonomous word structure (Qs) as strings. Autonomous word structure conditions may require an M^0 to split into two (or more) positions-of-exponence. Formally, this process is like that of Breaking in phonology: where language-specific ceilings on the complexity of a combination of features is reached in a given phonotactic environment, an atomic unit will split into a sequence of subunits.

Bleeding relations within this system arise in two ways. First, features are discharged only once, such that each feature may have at most one principal exponent. Discharged features then become visible to condition allomorphy of later affixes. Because a given feature may be discharged only once, a rule referring to a feature F will not apply to an input if F has been discharged-- even if F is discharged at a different position. This gives one kind of bleeding, feature bleeding. The second kind results from the fact that each obligatory position-of-exponence (Q) may be rewritten (discharged) only once as some string. Therefore, certain groups of affixes can be understood to 'compete' for being realized at this Q. This gives positional bleeding. When discontinuous bleeding happens, the featural and positional bleeding do not cooccur (as they must in a simple morpheme-based model).

Existing models fail to distinguish featural and positional bleeding. In the Extended Word Paradigm (EWP) Model of Anderson (1981, 1986, 1992), featural bleeding is generally limited to the domain of the disjunctive rule block, which is also the domain of that rule block's Elsewhere rule. These two domains do not coincide in Arabic, leading to paradoxes for this analysis. Special spurious identity rules or extra rules must be introduced to achieve the discontinuous bleeding generalizations. Anderson (1986) modifies the definition of disjunctive rule application to address a similar problem in Georgian. I showed that this modification is only partly successful for Arabic and again leads to ordering paradoxes.

The lexical affix model of (Lieber 1980, 1989, 1992) or Jensen (1990) suffers from being unable to directly capture the effects of affix hierarchies which EWP captures through rule ordering. Since this theory is committed conceptually to the idea that affixes combine freely to form words, subject to each affix's subcategorization frame, it has no way of ordering affixes in hierarchies of privilege when these compete for expression at single positions. The lexical model, however, unlike the paradigm model, permits both featural bleeding (by means of filling a "categorial signature") and positional bleeding (level-ordering, syntactic level of attachment). However, reducing positions to morphological strata leads to an ordering paradox in Arabic where certain forms must be prevented by appeal to Blocking within the lexical model. Blocking will not work at all within a syntax only approach such as Lieber (1992).

Finally, the abstract morpheme model of Halle (1989a, 1990, 1992), without further modification, ties features and positions of exponence together within the abstract morpheme. Bleeding of features and of positions is inseparable, unless abstract morphemes are permitted to split in various ways. Similar objections hold for the theory of Lumsden (1987, 1992), in which sets of affixes compete for realization at a single slot.

I proposed that the splitting of M^0s from syntax is accomplished by a series of rules which realize affixes at positions required by autonomous morphological structure. Each spell-out rule (1) places a feature or features in a position (2) discharges these features and (3) spells the position as a string.

In addition, I proposed the Feature Hierarchy Hypothesis, which in effect states that the same hierarchy of features determines the direction of Impoverishment of input morphosyntactic representations as determines the order of priority for affixes competing for a Q. This order of priority also obeys the Paninian principle identified with the Elsewhere Condition, but is here interpreted as a condition on rule ordering.

The reality of discontinuous bleeding was established through a detailed study of the historical development of the Afroasiatic prefix-conjugation. A number of analogical developments, including the extension of -ɣ in Tuareg and the distribution of *t-* in Northwest Semitic, Babylonian and S. Arabian, were shown to follow automatically in a theory where affixes actually discharge ("spell-out") features of the input, and where an elsewhere affix appears for any Q not rewritten by the end of the derivation. Certain other historical changes were shown to arise from the gain or loss of Impoverishment filters, such as the loss of gender in the plural Modern Arabic and the innovation of 1st person dual in S. Arabian, giving support to the theory of Impoverishment as developed in the Introduction.

CHAPTER 2: Towards a Theory of Person and Number Features

2.0 Introduction

In this chapter I detail a theory of person and number features. Any theory of person-number features must have three goals. The first is to predict what are the possible person-number categories in human language. The second goal is to show what person-number categories consist of: how they are represented and interpreted at the various levels of grammar. The third goal is to establish an evaluation metric according to which certain categories and inventories are understood as complex while others are simple from the point of view of language acquisition.

In chapter 1 and in the Introduction, I argued for a particular conception of certain kinds of morphology, which occurs after syntax to derive well-formed inflected words. In this component, rules of Morphology linearize syntactic constituents and convert abstract morphosyntactic representations to phonological strings. Similarly, I will be assuming an interpretive semantics, which subjects the output of syntax to rules of semantic interpretation. The model envisioned is as below:

(1) d-structure --------------------------------> Logical Form
(argument structure) ↓ (semantic interpretation)
Morphology
(phonological realization)

I will assume for simplicity that the same features figure in rules of phonological realization (Morphology) as figure in rules of semantic interpretation (Logical Form). This is not a necessary assumption. For example, one could assume that there are two sets of features, category features for Morphology, and semantic features at LF, and a universal principle mapping between these. I will not be pursuing this option, since, on my view, it entails an unnecessary complication.

The optimal theory of person-number features will by hypothesis allow straightforward interpretation at LF and realization at Morphology.

To give a simple example, assume that there is a person feature [1] and a number feature [plural]. In English at the morphological level, if these features are attached to a pronoun (D) which receives nominative case, this D will be realized phonologically as the string [wē] 'we.' At LF, [1] and [plural] condition rules of semantic interpretation requiring (1) that the speaker of the utterance have in mind himself or herself as one member of reference set of the argument associated with the pronoun, and (2) that the reference set is nonsingleton, i.e. contains at least two members.

Although this is a bit simplistic, the general picture is roughly of this form.

2.1 Person Features

2.1.1 Defining the Semantic Question

It has been usually assumed that the person categories are defined in terms of features such as [±speaker], [±hearer], as for example in Ingram's (1971/1978) cross-linguistic study of the person categories. In this section, I present the assumptions about semantic interpretation that underlie the postulation of such primitive features and not others.

The interpretation of person features must recognize certain *discourse roles* as primitives: these roles are among the deictic markers of a speech-act, situating the speech-act with respect to its place, time, and, in the case of person features, its participants. The primary distinction is between participants in the speech-act and nonparticipants, what Hockett (1966) refers to as "local" vs. "nonlocal" arguments. The speaker and hearer are local participants, while other parties, neither speaker nor hearer, are nonlocal nonparticipants:

(2) participants: speaker or hearer
nonparticipants: neither speaker nor hearer

The participant division can be subdivided into speakers and hearers.[1] I will use numerals to designate these subroles:

(3) 1 = speaker
2 = hearer
3 = not speaker, not hearer

I will be assuming that all syntactic arguments bear anaphoric indices of the type proposed in Chomsky 1981 and Lasnik 1981. Such indices are applied freely to arguments subject to the condition that arguments bearing the same indices be coreferent. (Inversely, "accidentally" coreferent arguments need not bear the same indices). What is important for

[1] I ignore for now questions of the proximate vs. obviative distinction among third person arguments.

the theory of person number features is that discourse roles are assigned to anaphoric indices and not to real world referents.

To understand this consider:

(4) Cicero$_{\{i\}}$ and Tully$_{\{j\}}$, why, they$_{\{i,j\}}$ are just one guy$_{\{k\}}$!

The names *Cicero* and *Tully* refer to the same real world individual, but, from the point of view of anaphoric indexation, they bear distinct indices {i} and {j}. For this reason, the plural pronoun *they* may be used to refer to the two *concepts* of the same real world individual. It is to such concepts, as identified by anaphoric indices, that discourse roles are assigned.

It is therefore possible, for any given sentence, to set up an assignment of discourse roles to anaphoric indices in the manner of a function from indices to roles. For example, sentence (4) could have the assignment

(5) $\{ \langle i,3\rangle, \langle j,3\rangle, \langle k,3\rangle \}$

Any sentence will have such an assignment of discourse roles, and I will assume that such an assignment is free.[2] For expository convenience, I will write N for the set of roles to which the set of indices of a given argument is paired by this assignment. For example, all three arguments in (4) will have N = {3}.

2.1.2 Natural Language Person Syncretisms

Let us now consider some examples of such sets N:

(6)
a. {1}
b. {2}
c. {3}
d. {1, 2}
e. {1, 3}
f. {2, 3}
g. {1, 2, 3}

[2] Note that no anaphoric index may be paired with more than one role.

There are obviously seven distinct sets which may be formed from the three discourse roles. Less obvious is that natural languages do not employ a seven-way distinction, but instead have a maximal division of four person categories, as shown in (7):

(7)	a. {1} or {1, 3}	1st person (exclusive)
	b. {1, 2} or {1, 2, 3}	1st person (inclusive)
	c. {2} or {2, 3}	2nd person
	d. {3}	3rd person

Although certain languages permit compound pronominal forms which demand a greater subdivision of person types, I will show how these may all be understood as arising from the incorporation of one pronoun type in (7) into another.

There is no "exclusive/inclusive" distinction for the 2nd person: (6c) and (6f) invariably fall together. In other words, there is no language with simplex pronouns distinguishing {2} from {2,3}, that is, an argument picking out only hearers to the exclusion of non-hearers, {2} only.

Greenberg (1988:14) sums up this universal:

> [there is] a universal principle, with a few rare and doubtful exceptions, namely that languages do not distinguish in what are usually called second person nonsingulars between 2 and 3, as long as the sets (2,2), (2,2,2) etc. are present depending on the appropriate number category.

By (2,2) Greenberg means (in effect) two indices assigned the role 2, i.e. something like { <x,2>, <y,2> }.

Similarly, there is no distinction between {1, 2} and {1, 2, 3}. No 1st person inclusive {1,2} argument necessarily excludes {3}. For example, in English the pronoun *we* can never in principle exclude a party not addressed. For example, if I say, "we learn more about morphology every day," there must always be in principle an interpretation in which 'we' may include any third party. The same is true in all languages. I will later show that an alleged counterexample from Sierra Popoluca (Foster & Foster 1948, Zwicky 1977), has been misanalyzed.

Finally, no language has a distinction between {1} and {1,3}. This fact becomes obvious in the plural: no language has a 1st person plural argument with only multiple speakers, i.e. just the role {1}.

The task at hand is to give a set of person features which permits all and only those syncretisms which are observed. To this end, consider how the four person categories may be represented using parentheses for the optional role(s):

(8) a. {1, (3) } 1st person (exclusive)
b. {1, 2, (3) } 1st person (inclusive)
c. {2, (3) } 2nd person
d. {3} 3rd person

The following generalization emerges:

(9) Generalization I:
No person category specifically excludes {3}.

Of course, in the singular, 1st person exclusive and 2nd person exclude {3}, but only indirectly, in virtue of having only one index and therefore being unable to have more than one role. Because no category excludes {3}, there can be no feature [-3] which could enforce this exclusion. Categories in (6) rendered impossible by Generalization I include: {1}, {2}, and {1, 2}, since each of these excludes {3}.

Additionally, another pattern can be discerned in (8):

(10) Generalization II:
Unless {3} is the only permitted role (= 3rd person), {3} is optional

In other words, except for the 3rd person, no category specifically *requires* the role {3}. Recall now that there is a major bifurcation of discourse roles into participants {1,2} and non-participants {3}. The generalization can now be restated:

(11) Generalization II (Revised):
Only an argument which excludes participants can require {3}

Another way to understand this is that the 3rd person can be viewed in two ways. It can be defined as that category which is specifically nonparticipant (i.e. Benveniste 1956/1972: "the non-person") or as that category which specifically requires {3}, that is, the [+3] category.

Clearly, a feature such as [+3] is too powerful, since, other things being equal, we should expect a distinction between {2} and { 2, 3}, with the latter specifically demanding a third person. Instead, I will follow Farkas (1990) in assuming that the 3rd person is defined as [-participant], i.e. that category excluding the roles {1, 2}. Since {3} is the only remaining role, a [-participant] category is *de facto* a {3} category.

Categories in (6) which are rendered impossible by Generalization II are: {1, 3}, {2, 3}, and {1, 2, 3}. Each of these requires {3} but does not exclude {1} or {2}.

Generalizations I and II also suffice to exclude a number of categories which can be logically composed from the sets in (6). These are given below:

(12)

	Set	Generalizations Flouted I	II
a.	{ (1), 2}	*	
b.	{ (1), 3}		*
c.	{ (2), 3}		*
d.	{ 1, (2)}	*	
e.	{ (1), 2, 3}		*
f.	{1, (2), 3}		*
g.	{(1), (2), 3}		*

As can be seen above, all the pseudo-categories which flout Generalization I exclude {3}; the ones which flout Generalization II require {3} while permitting {1} or {2} or both.

I follow Ingram (1971/1978) and Hale (1973) in postulating the features [±I] and [±you] or their notational equivalents to define the remaining permissible categories as below:

(13)

[±I]	[±you]	[±participant]	sets	name
+	-	+	{1, (3) }	1st excl
+	+	+	{1, 2, (3) }	1st incl
-	+	+	{2, (3) }	2nd
-	-	-	{3}	3rd

The feature [+I] requires {1}; the feature [-I] excludes {1} and likewise for [±you] for the {2} role. The feature [-participant] excludes {1} and {2}; the feature [+participant] requires *either* {1} or {2} (but not necessarily both). Note that the value [+participant] does not exclude {3}-- it merely requires {1} or {2}. In this way, no category can be made to specifically exclude {3}.

I will assume that the features above condition the appropriate rules of semantic interpretation at LF via the assignment of discourse roles to anaphoric indices at a level of discourse representation.

Given that there are only four categories in (13), the minimal number of features which can differentiate these four is two. The value of the feature [±participant] is always redundant if the values for [±I] and [±you] are known. In the next section I review evidence for why the value [+participant] is a natural class in morphology.

2.1.3 The Value [+participant]

A number of morphological rules refer to the [+participant] categories as a natural class and therefore provide evidence that [+participant] is a feature visible at Morphology. I will cite several of these examples.

The first case comes from Winnebago (Lipkind 1945, Ken Hale, p.c.). The Winnebago free personal pronouns distinguish only [±participant]:

(14)	nee	'I' or 'you'	[+participant]
	ʔee	'he/she'	[-participant]

Disambiguation of 'I' and 'you' is accomplished through verbal morphology. It is important to note that neither the set notation nor a feature system without [±participant] can capture the class of {1} *or* {2}. Similarly in Winnebago, the augmented number suffix is -wi for 1st and 2nd person subjects, but -ire for third person subjects. Again the morphological split is between the [+participant] categories (-wi) and the [-participant] ones (-ire).

The second case comes from Navajo. 1st and 2nd person plural pronouns are homophonous (Ken Hale, p.c.):

(15)	a. nihí	'we / you (pl.)'	independent
	b. -nihi-	'we / you (pl.)'	object clitic (toneless)

Finally, 1st and 2nd person accusative agreement affixes are homophonous in Lummi, a Salish language of British Columbia (Jelinek 1992). Consider the following examples:[3]

[3]CT = control transitivizer, a suffix marking valence and agentivity. The sentences imply that the agent hit the patient on purpose rather than by accident.

(16) a. t'əm'-t-óŋəs = sxʷ
hit-CT-1sg ACC=2sg NOM

'You hit me.'

b. t'əm'-t-óŋəs = sən
hit-CT-2sg ACC=1sg NOM

'I hit you.'

c. t'əm'-t-óŋəɬ = sxʷ
hit-CT-1pl ACC=2sg NOM

'You hit us.'

In (16a) and (16b), the same affix -oŋəs realizes a 1sg object agreement or a 2sg object agreement. In (16c) a distinct form -oŋəɬ realizes a 1pl object. Jelinek reports that 2nd plural is identical to 2 sg as well. The realizations of the accusative agreement affixes must be:

(17) a. [+participant +I -sg] oŋəɬ
b. [+participant] oŋəs

Rule/affix (17a) will be realized first since it is more specific. The class of categories which are realized as oŋəs form an elsewhere class defined by [+participant]. This provides more evidence that [+participant] is a morphologically active feature in at least some languages.

2.1.4 A Universal Hierarchy of Persons?

Not all languages have the four way distinction permitted in (13). In particular, the inclusive/exclusive distinction is absent from Indo-European and Semitic, among other language families. The first person arguments in these languages have the discourse roles {1, (2), (3)}. The three persons, in say, English, are as below:

(18)	{1, (2), (3)}	1st person (generic)
	{2, (3) }	2nd person
	{3}	3rd person

As Zwicky (1977) observes, it is significant that no language has a functional syncretism of the sort:

(19) *{ (1), 2, (3) } "syou"

"Syou" requires at least one addressee, and may include any others, including the speaker. Nothing said so far will exclude this category, which is the syncretism of 1st inclusive and 2nd person. (I delay discussion of the alleged "syou" of Algonquian until the next section). "Syou" passes by Generalization I since it doesn't exclude {3}, nor does it require {3}, so it passes Generalization II. Why shouldn't "syou" be a possible category?

Compare the generic 1st person as in English with "syou":

(20) {1, (2), (3)} 1st person (generic) "we"
{ (1), 2, (3)} "syou"

"Syou" is like "we" except that, rather than having {2} optional, {1} is optional. In order to exclude "syou," we must appeal to a third Generalization:

(21) Generalization III:
The (1) role can never be optional.

In other words, a category either necessarily includes the speaker (1), or necessarily excludes the speaker. Generalization III will also exclude (12a,b,g,e) as well as "syou."

Although he presents the matter in rather different terms, Zwicky (1977) suggests that there exists a Universal Hierarchy of Persons of the form

(22) Universal Hierarchy of Persons (UHP):
$1 > 2 > 3$

In the terms presented here, this hierarchy has the following correlate. We have already observed that the {3} role is optional unless it is the only role required. Now consider the {2} role. Observe that the {2} role is optional only in case the {1} role is obligatory (in the generic 1st person). In such a case, the {3} role is also optional. Finally, the {1} role is never optional. We can make the following generalization:

(23) Generalization IV:
A role may be optional only if a higher role on the hierarchy is required.

Now that we have these descriptive generalizations the question becomes how to make them follow from the theory of person features. Recall that Generalizations I and II follow from the absence of [±3]. Some aspect of the theory must now permit "we" while barring "syou."

Consider these two categories in feature terms:

(24) [+I αyou] 'we'
[αI +you] 'syou'

The 'we' category requires the role {1} and optionally includes {2}. This translates to any value for [±you], either [+you] -- including the addressee -- or [-you] -- excluding the addressee-- is fine. The inverse situation holds of 'syou' with (1) being the optional role.

How do such "underspecified" categories come about? I argued in the Introduction that fully specified morphosyntactic categories are Impoverished at the level of Morphology by language-specific filters. The filter which gives rise to the syncretism of 1st inclusive and exclusive as a generic 1st person category is:

(25) *[+I αyou]

Recall that by the Feature Hierarchy Hypothesis (section 1.10), filters must be combined with feature hierarchies to determine which feature will be deleted in an offending combination. Here I postulate a simple hierarchy:

(26) I > you

When encountering a matrix containing a value for both [+I] and some value for [±you], the filter (25) combined with the hierarchy (26) will delete the value for [±you]. Thus, 1st inclusive and 1st exclusive neutralize to a generic [+I] category.

Because of the hierarchy of features (26), a neutralization of the type "syou" cannot come about through Impoverishment. This is because the "syou" category can be derived only by deleting [±I] in the presence of [+you] with a filter of the form *[αI +you]. This deletion cannot happen since such deletions always obey the hierarchy of features.

Returning now to the Universal Hierarchy of persons, it is clear that it is not wholly necessary. The conditions on the appearance of role (3) follow from the definition of

[±participant], which captures both Generalizations I and II, and need not be separately stipulated. This leaves only the mini-hierarchy among participant features in (26).

Not all languages have the filter (25). Following the ideas of Calabrese (1988, 1992) as presented in the Introduction. I will assume that the filter in (25) as well as the various person features and their semantic interpretations and implicational relations are part of Universal Grammar. When positive evidence presented to a child shows that 1st inclusive and 1st exclusive are distinct categories, then the filter (25) is "switched off." In section 2.1.7 I return to other filters of the person system, showing how a hierarchy among these can derive an evaluation of complexity for person inventories.

2.1.5 The Putative 'syou' of Algonquian

There has been considerable puzzlement in the literature surrounding a putative case of 'syou' in the Algonquian languages (Zwicky 1977, Jolley 1983). In the verbal and nominal morphology, both 1st person inclusive and 2nd person take the clitic (Proto-Algonquian) *k- or its reflex in the various Algonquian languages, while 1st exclusive takes *n-. One might then be led to suppose that *k- represents 'syou.'

However, this is not a systematic neutralization at the level of Impoverishment. In other words, the 1st inclusive and 2nd person categories continue to be distinct in the Morphological component, but *happen* to take the same clitic prefix. These two categories are distinguished by suffixal morphology and are not fully neutralized. Consider the examples from Plains Cree (Bellegarde & Ratt 1989:38-39):

(27) a. ni-miciso-nān
1CL - eat- 1 pl indic

'We (excl.) eat'

b. ki-miciso-nāwāw
2CL - eat- 2 pl indic

'You (pl) eat'

c. ki-miciso-naw
2CL - eat - 21 indic

'We (incl.) eat.'

The prefixed clitics have the following forms:[4]

(28) a. [+I -you] ni (t)-
b. [+you] ki (t)-

Although rule (28b) collects the class of [+you] categories, 1st inclusive and 2nd person, these remain distinct in Morphology, because 1st person inclusive and 2nd person have distinct agreement suffixes -- -nāwāw for 2nd plural and -naw for 1st inclusive.[5] The [±I] feature distinguishing them must not be deleted by Impoverishment. The syncretism of 1st exclusive and 2nd is a property of the rules which realize the shape of the clitic prefix. Spell-out rules of this sort may refer to any natural class defined by a feature, here [+you]. This does NOT mean that Algonquian has a *category* 'syou.' To have such a morphological category, in the terms presented here, is to have a rule of Impoverishment which combines 1st inclusive and 2nd person. Algonquian has no such rule: the neutralization of the value [±I] is only at the level of spell-out (realization of k-).

One reason why a solution of this sort was (implicitly) deemed objectionable by Zwicky (1977), for example, is that it requires reference to the value [-you] in rule (28a). Zwicky argues that morphological realization rules -- whether of the sort which realize phonological material (spell-out) or of the sort which sequence this material (templates, M^0 splitting operations) -- do not refer to the negative values [-I] or [-you]. It was accordingly understood as desirable to eliminate these values from morphological representation totally. The five possible categories can then be represented as below:

---

[4]The final *t* of the prefix occurs only before vowels.
[5]Because these two suffixes, -nāwāw '2 pl' and -naw '1 incl pl' share some phonological material, one might suppose that they are amenable to further analysis. While in the indicative forms presented here the 2 pl suffix looks more like the 1 incl suffix than the 1 pl suffix -nān, in the subjunctive the opposite holds:

Plains Cree agreement affixes (partial) (Bellegarde & Ratt 1989:39)

	indic.	subjunctive
1 excl pl	-nān	-yāhk
1 incl pl	-naw	-yahk
2 pl	-nāwāw	-yēk

Both vowel length and vowel height differentiate these suffixes. Although similar in form, each is probably atomic at the synchronic level.

(29)	[+I]	[+you]	Name
	+	Ø	1st person
	+	+	1st person inclusive
	Ø	+	2nd person
	Ø	Ø	3rd person

Unfortunately, such a scheme, partly implicit in Zwicky's (1977) conception of morphosyntactic category features, is too ambitious. While it is generally true that morphological rules "see" only the [+] values for the features [±I] and [±you], I now consider evidence that shows that, at least *some of the time*, the values [-I] and [-you] must remain visible in morphological representation.

2.1.6 The Person Affixes of Mam

Mam is a Mayan language spoken in northwest Guatemala. Mam provides rather striking evidence that both values of [±I] and [±you] must be visible in Morphology. The system of person affixes for the verb is given below (data from England 1983:56):

(30)		Set A	Set B	Enclitic
	1 sg	n- / w-	chin-	-(y)a
	2 sg	t-	Ø / tz- / tz'-/ k-	-(y)a
	3 sg	t-	Ø / tz- / tz'-/ k-	
	1 excl pl	q-	qo-	-(y)a
	1 incl pl	q-	qo-	
	2 pl	ky-	chi-	-(y)a
	3 pl	ky-	chi-	

The set A prefixes are ergative and the set B prefixes are absolutive, and the variants within each set, e.g. n- vs. w- for 1 sg set A, are conditioned by phonology and by the mood of the verb (England 1983:56). An example verb is given below:

(31) a. ma tz'- ok t-tzeeq'an-a
recent.past 3sgB directional 2sgA.hit-ENCL

'You hit him/her/it'

The absolutive (transitive patient) appears as tz'- followed by a directional (adverbial element), followed by the ergative t-, the verb, and finally the enclitic -a, whose distribution I focus on later.

As is typical of Mayan (and also many languages with ergative morphology such as Inuit), the same set of person markers that is used for ergative agreement (set A), also marks the possessor of a nominal, as in the following examples (England 1976:261):

(32)		
	a. n -wi:xh -a	'my cat'
	b. t -wi:xh -a	'your (sg) cat'
	c. t -wi:xh	'his/her cat'
	d. q -wi:xh -a	'our (excl) cat'
	e. q -wi:xh	'our (incl) cat'
	f. ky -wi:xh -a	'your (pl) cat'
	g. ky -wi:xh	'their cat'

There is an isomorphy between the set A prefixes and the set B prefixes in (30), so any morphological analysis that can derive one pattern can derive the other. For simplicity, consider below the set A (ergative) forms, which can be tabulated as:

(33)		singular	nonsingular	
	[+I -you]	n- / w- -a	q- -a	1 excl
	[+I +you]	*	q-	1 incl
	[-I +you]	t- -a	ky- -a	2
	[-I -you]	t-	ky-	3

Two interesting patterns can be discerned right away. The first is that the division in the prefixes is between the 1st person categories and all others. As England points out, the choice of prefix is sensitive only to the feature [±I] and to number. The following rules show the correct distribution:

(34)		Set A	Set B
	a. +I −sg	q-	qo-
	b. +I	n- / w-	chin-
	c. −I −sg	ky-	ky-
	d. −I	t-	Ø / tz- / tz'-/ k-

This pattern of distribution is not an unassailable argument in favor of [-I] being visible in Morphology. Because the rules in (34) are in a bleeding relation, they can be rewritten:

(35)

		Set A	Set B
	a. +I −sg	q-	qo-
	b. +I	n- / w-	chin-
	c. −sg	ky-	ky-
	d. Elsewhere	t-	Ø / tz- / tz'-/ k-

In other words, if it is insisted that [-I] is invisible at Morphology, rules may still refer to the [-I] class as the Elsewhere class, provided earlier rules discharge [+I].

Stronger evidence in favor of the visibility of [-I] and [-you] comes from the second interesting pattern, the distribution of the enclitic suffix -a. The enclitic -a occurs in the 1st person singular, 1st person exclusive plural and 2nd person singular and plural. These categories are either [+I -you] OR [-I +you]. These are all and only those categories where the values for [±I] and [±you] *disagree*. The only way to express these categories as a natural class is by a rule using the α–notation (recall that parentheses indicate a previously discharged feature, not an optional feature):

(36)

	a. (αI) −αyou	-a
	b. Elsewhere	-Ø

In other words, -a has the value of [±you] which is opposite of the value for [±I] as discharged in the prefix. For rule (36a) to work properly, both the [+] and [-] values of [±I] and [±you] must be visible at Morphology.

By creating a special -Ø suffixation rule for the 1st person inclusive, it is possible to avoid using the [-] values. Two such analyses are given below:[6]

(37)

	a. (+I) +you	-Ø
	b. +I	-a
	c. +you	-a
	d. Elsewhere	-Ø

[6]England (1976) shows that, in all likelihood, the *-a* enclitic was originally a politeness marker, suffixed to the third person form to give an honorific 'you.' Cf. Italian *Lei* / *Loro* meaning 'you (sg/pl)' but inflected as third person pronouns.

(38)	a. (+I) +you	-Ø
	b. +participant	-a
	c. Elsewhere	-Ø

The two rules (37b) and (37c) can be collapsed as the [+participant] categories remaining after 1st incl is discharged by the [+I +you] rule. Thus the two rules can be collapsed as in (38b).

In the above rules, the elsewhere class is *de facto* [-I -you]. Thus, this analysis disguises the fact that the distribution of -Ø is in fact the cases where the features of [±I] and [±you] *agree*. (37) or (38) could just as easily be written:

(39)	a. (αI) αyou	-Ø
	b. Elsewhere	-a

To summarize, analyses which permit both the [-] values of [±I] and [±you] to be visible at Morphology and also permit α-notation ((36) or (39)) can straightforwardly express the distribution of enclitic -a vs. -Ø in Mam. Alternative analyses must assume (at best) that there are either two homophonous -Ø affixes, one especially occurring in the 1st inclusive (38).

There are two arguments against the homophony analyses (37) and (38). The notion underlying both arguments is that the forms with enclitic -a and those without enclitic -a form two natural classes with respect to other morphological phenomena. The homophony analysis cannot express these generalizations. These two arguments are presented in the next two subsections.

2.1.6.1 Emphatic Possessor

The first phenomenon, discussed in England (1976:260-61), is one type of emphatic possessor morphology. One way to express an emphatic possessor in Mam is to double the possessor prefix on the noun. This results in 'simple contrastive emphasis' (England 1983:143). Interestingly, however, only those forms which when unemphatic have the enclitic -a may form such emphatics. Final enclitics do not appear with the doubled-prefix emphatic (England 1976:261):

(40)	Unemphatic		Emphatic	
	a. n-wi:xh-a	'my cat'	w-n-wi:xh	'*my* cat'
	b. t-wi:xh-a	'your (sg) cat'	t-t-wi:xh	'*your* (sg) cat'
	c. t-wi:xh-Ø	'his/her cat'	*	
	d. q-wi:xh-a	'our (excl.) cat'	q-q-wi:xh	'*our* (excl) cat'
	e. q-wi:xh-Ø	'our (incl.) cat'	*	
	f. ky-wi:xh-a	'your (pl.) cat'	ky-ky-wi:xh	'*your* (pl) cat'
	g. ky-wi:xh-Ø	'their cat'	*	

Observe that wherever the unemphatic possessed form has a zero-suffix, the emphatic cannot be formed. Wherever the unemphatic has the enclitic -a, the emphatic is well-formed, but lacks the enclitic, instead doubling the prefix. Also note that the doubling of the prefix is morphological, rather than simply a phonological reduplication, as is shown by the 1sg form w-n-wi:xh '*my* cat,' where the first prefix is a different allomorph of the 1sg than the second prefix.

According to England (1983:143), the w- allomorph of [+I] is conditioned by an underlying vowel, which is 'usually not pronounced due to vowel dropping rules'; it is possible that this vowel is the enclitic -a appearing before the noun rather than after:

(41) ?? w-a-n-wi:xh --> w-n-wi:xh '*my* cat'

At the very least, the pattern in (40) shows that the forms with enclitic and without form two natural classes with respect to some rule of morphology. This militates against any analysis that treats the various instances of -a as homophonies.

2.1.6.2 Enclitics in Patient-Agent Agreement

The second phenomenon showing that the forms in -a behave as a natural class as opposed the forms in -Ø comes from patient-agent agreement. As shown in example (31), a transitive verb inflects for both an agent argument and a patient argument, with the patient argument realized first in the word as a set B prefix and the agent argument to the left of the stem as a set A prefix. This is uncomplicated. The complication arises in the distribution of the enclitics in transitive verbs. To quote England (1983:61):

> Although there are two sets of prefixes to indicate the agent and patient of a transitive verb, there is only one set of enclitics, which refers to either the agent or both the agent and the patient. If the agent does not require an enclitic, then the patient cannot have an enclitic. If the agent does require an enclitic, then the patient may or may not be indicated by the enclitic . . . Disambiguation is possible through specification of third person arguments in noun phrases and through use of the antipassive.

Sense can be made of this pattern of facts once it is understood that the -Ø suffix may be the marked suffix and -a is merely the elsewhere realization of the enclitic position of exponence.

I will assume that the realization of the enclitic position occurs on two successive cycles, as determined by the order of adjunction of the Patient and Agent agreement in syntax. The syntactic structure then must be mapped onto an autonomous word structure which includes the enclitic position:

(42) $[\mathrm{AGR}_{\mathrm{Patient}}\ [\ \mathrm{AGR}_{\mathrm{Agent}}\ \mathrm{VERB}]\ ]$

$\mathrm{Q}_{\mathrm{Patient}}\ \mathrm{Q}_{\mathrm{Agent}}\ \mathrm{VERB}\ \mathrm{Q}_{\mathrm{Enclitic}}$

By hypothesis, the enclitic position realizes the AGR feature [±you]. However, as I will show, because the rules operate cyclically, the enclitic position will be filled either by -Ø or by -a by the end of the first cycle in any event. Thus only $\mathrm{AGR}_{\mathrm{Agent}}$ splits into two positions of exponence as shown above. This presents a case of M^0 splitting analogous to that of the Arabic INFL split discussed in Chapter 1.

To clarify what is going on, in (43) I exhibit the four logical patterns of occurrence, with respect to the presence or absence of -a in the agent and patient. 'A' stands for the ergative set A prefix, 'B' for the absolutive prefix. The form which is ruled out by England's generalization is (43c):

(43)	a.		A	Stem	-a	
		B			-a	both have enclitic
	b.		A	Stem	-a	
		B			-Ø	agent has enclitic, patient doesn't
	*c.		A	Stem	-Ø	
		B			-a	patient has enclitic, agent doesn't
	d.		A	Stem	-Ø	
		B			-Ø	neither has enclitic

Examples of the four types are given below, with the tense and directional markers omitted for clarity:

(44)	a. . . . tz' ... n-tzeeq'an-a	'I hit you (sg).'
	b. . . . tz' ... n-tzeeq'an-a	'I hit him/her/it'
	c. *. . .qo ... t-tzeeq'an-a	* 'he hit us (excl.)'
	d. . . . qo ... t-tzeeq'an	'he hit us (incl.)'

When both the agent and the patient have the enclitic (44a), the form does not surface with a long *-a-a*. This shows that there is only one position of exponence for the enclitic. Thus a form with *-a* is ambiguous between the reading in which the agent takes *-a* (43b)/(44b) and the reading where there would be two enclitics (if there were two positions to realize them): (43a)/(44a). Observe that (44a) and (44b) are homophonous.

In the schema in (43), the fact that there is but one position for the enclitic realizations is shown by putting the two enclitic realizations in a column in each form. The first line in each form is meant to represent schematically those elements which are realized on the first cycle, having been conditioned by the agent features; the second line is the second cycle of patient features.

Consider again the enclitic rules:

(45)	a. (αI) αyou	-Ø	= (39)
	b. Elsewhere	-a	

These rules apply on the first cycle, the agent cycle, and discharge the enclitic position, either by filling it with -Ø or by filling it with the default suffix -a. Even if the rules were to apply again on the second cycle, they could not, since the enclitic position has been filled/discharged. Therefore, whatever enclitic actually appears reflects the agent features from the first cycle.

Observe now that any analysis which does not somehow incorporate the notion of a zero-affix in some way will be unable to explain why the forms in (43c) could not have the expected enclitic -a conditioned by patient features on the second cycle. In a paradigm model, even an identity rule on the earlier cycle (giving the effect of a -Ø affix) could not be made disjunctive with the rules of the second cycle.

It will be instructive to consider derivations of the four types in (43):

(46)

Input:	a. [-I +you [+I -you [VERB]]		b. [-I -you [+I -you [VERB]]	
Cycle I				
(35b)		n- VERB		n- VERB
(45b)		n- VERB -a		n- VERB -a
Cycle II				
(35d)	tz'-	n- VERB -a	tz'-	n- VERB -a
(45a)			cannot apply since suffix filled	
(45b)	cannot apply since suffix is filled			
Input:	c. [+I -you -sg [-I -you [VERB]]		d. [+I +you -sg [-I -you [VERB]]	
Cycle I				
(35d)		t- VERB		t- VERB
(45a)		t- VERB -Ø		t- VERB -Ø
Cycle II				
(35a)	*qo-	t- VERB -Ø	qo-	t- VERB -Ø
(45a)			cannot apply since suffix is filled	
(45b)	cannot apply since suffix is filled			

The bad form (44c) is shown in derivation (46c) above. When the patient is expected to take the enclitic -a but the agent takes -Ø, then the patient enclitic cannot appear, having been pre-empted by the -Ø of the agent. From such an input, only a zero-suffixed form may be produced.

2.1.6.3 Processing and Ambiguity

There is a curious asymmetry in the pattern of acceptability in (46). The generalization is that the only bad form is the form in which enclitic *-a* *would have been* supplied on the second cycle were it not for the pre-empting -Ø of the first cycle (43c), (44c), (46c). Inversely, however, the form is fine in which enclitic -Ø cannot appear on the second cycle owing to having been pre-empted by *-a* on the first cycle. Thus, (47a) is ambiguous, but (47b) is not:

(47) a. . . . tz' ... n-tzeeq'an-a
Ambiguous: 'I hit you (sg)' or 'I hit him/her/it'

b. . . . qo ... t-tzeeq'an
Not ambiguous: means only 'he hit us (incl.)', not 'he hit us (excl.)'

In fact, however, some pairs like (47b) are marginally ambiguous, with the form where the patient is expected to have an enclitic having a (?) judgement. England (1983:62) cites an ambiguity between 'he/she/it hit you.' (? judgement) and fully acceptable 'he/she/it hit him/her/it.' Apparently some speakers accept the (?) forms while others reject them.

My hypothesis regarding these variable judgements is that it is crucial here to separate morphological facts from processing facts. Consider (47b) for example. When this form is processed, the hearer is inclined to assume that what appears in the enclitic position (here -Ø) is what would appear here for both agent and patient agreement if realized separately. In order to interpret the form as 'he hit us (excl.)' the form must be processed such that there is a 'ghost' enclitic *-a* which could not be realized because of the pre-empting -Ø. Therefore, in processing such forms, the hearer will provide only the most natural interpretation, in which there is no 'ghost' *-a*. The underlying assumption in processing would seem to be 'what you see is what you get.'

Verb forms like (47b) are like garden path sentences from the point of view of processing. When the speaker hears the -Ø enclitic, i.e. the phonetic absence of a clitic, then he or she is led down the garden path of supposing that the -Ø does not conceal a ghost enclitic.

It is then not surprising that among the forms which England found to be marginally ambiguous are the less complicated category combinations, such as 3sg --> 2sg vs. 3sg --> 3sg. More complicated ambiguities such as 3 --> 1 incl vs. 3 --> 1 excl are both rarer in occurrence and less likely to impede communication. This is consistent with the hypothesis that the forms barred by England's generalization are not, strictly speaking, ungrammatical, but rather, unprocessable.

It should be borne in mind, however, that for any given combination, rules of morphology will generate *some* form. The 'impossible' combinations such as 3 --> 1 excl in fact do generate forms with a -Ø suffix. A form with an overt enclitic cannot be generated. The problem for these combinations is that the form so generated is a *morphological garden path.* Speakers with a view to communicating effectively naturally avoid such forms and use the antipassive instead.

If such forms are really garden path words whose ill-formedness should be attributed to a processing problem rather than to a grammatical one, we should expect to find cases in which processing difficulties should be overridden by context. Although I have found no such examples in England's description of the verbal facts (aside from the variability of speaker judgments in some forms), there is one piece of evidence that may support this idea.

In addition to the emphatic possessor morphology described in 2.1.6.1, there is another type of emphatic possession, which combines both emphasis (focus) and topicalization. This type of emphatic possessor is formed by preposing and postposing the person marker and adding the enclitic *-a* (England 1983:143-144). The following examples contrast the two types of emphatic possessed nouns:

(48) a. w-n-jaa kyee?yex Ø-kii-n
[+I]-[+I]-house nice [-I]-see-antipassive

'*My* house looks nice.'

b. n-jaa-w-a aj jaa kyee?yex Ø-kii-n
[+I]-house-[+I]-encl relative house nice [-I]-see-antipassive

'It is *my* house that looks nice'

In (48a) there are two preposed person markers, giving the untopicalized emphatic; in (48b) the noun is sandwiched between the two person markers. Because the topicalized form occurs in a particular syntactic context, it is reasonable to suppose that some syntactic process is involved in constructing the topicalized emphatic noun-word; but I have at present no analysis of this.

What is significant about the topicalized emphatic is that, like the untopicalized emphatic, it occurs only in the person categories where the enclitic has the realization *-a*, namely 1 excl and 2, but unlike the untopicalized emphatic, the presence of an overt noun phrase allows a 3rd person interpretation as well:

(49) a. t-jaa-t-a
[-I]-house-[-I]-encl '*your* house'

b. t-jaa-t-a xu?j
[-I]-house-[-I]-encl woman 'the *woman's* house'

This example is intended to show that the presence of an overt NP possessor is sufficient to cancel the [+you] interpretation afforded by the enclitic *-a* in (49b). We must assume that the enclitic in these forms is realized obligatorily without being conditioned by the feature [+you]. A fuller understanding of the syntax of Mam topicalization and the role of the enclitic therein is required firm up this suggestion.

To summarize the discussion, Mam provides evidence for a number of theoretical issues. The first is the visibility of [-I] and [-you]. Because the distribution of the enclitic marker can be accounted for only through reference to these values, then these values must be present in Morphology, at least some of the time. Second, the pattern of markedness in the emphatic possessor forms reveals that among the enclitics -a and -Ø, the zero-form is the marked one and -a is realized as the elsewhere case. This provides evidence in favor of rules which explicitly realize positions of exponence as phonetic nulls. Third, the distribution of enclitics in verbs agreeing with two arguments shows that the realization rules operate cyclically, here with the agent features visible first. More specifically, because the enclitic position will be filled/discharged at the end of the first cycle, patient features visible on the second cycle never appear in the enclitic position. A -Ø realized in the enclitic position on the first cycle pre-empts an -a which would appear there on the second cycle. Finally, I attributed the unacceptability of certain forms in which -a is pre-empted by -Ø to a processing problem. These forms, being ambiguous with other forms in which -Ø is pre-empted by -Ø, are really garden path *words*. As such, they are still acceptable for some speakers for certain combinations of arguments (which I suggest are more commonly occurring).

2.1.7 A Hierarchy of Person Feature Filters

On the basis of the Mam data, we are forced to assume that [-I] and [-you] can be visible in Morphology. An obvious question now is: *must* these values be visible?

Suppose that [-I] and [-you] are visible at Morphology in every language, universally. If this is so, then there is no way immediate way to explain why these values are so rarely required in Morphology. More specifically, the classes defined by these features, namely 2nd and 3rd person [-I] and 1st and 3rd person [-you], are almost never a natural class with respect to morphological rules, as Zwicky (1977) points out. If both values are visible at Morphology, then in terms of a simple feature counting evaluation metric for rule complexity, a [-I] rule should be no more complex or unusual than a [+I] rule. Yet *every* language has a [+I] rule which defines its 1st person category (a universal category). Almost *no* languages have [-I] rules. Any explanatory theory of Morphology must somehow express this fact about language in a direct way.

One approach is to assign some cost value to morphological rules depending upon what features they refer to. Any rule referring to [-I] would be highly costly, whereas a rule referring to [+I] would be nearly cost-free. I will not be pursuing this approach to complexity, partly because an evaluation of cost in these terms is too vague. Instead, I propose that cost be encoded through hierarchies of filters of the kind proposed by Calabrese (1988, 1992) for phonology.

Recall from the Introduction that Calabrese encodes the complexity of certain segment types by means of a hierarchy of filters which also encode implications regarding inventories of segments. For example, it has been observed that a [±ATR] contrast in low vowels is exceedingly rare cross-linguistically, and furthermore, any language with such a contrast also has a [±ATR] contrast for both high and mid vowels. Any language with the [±ATR] contrast in high vowels has this contrast for mid vowels. Thus Calabrese postulates the hierarchy (50):

(50)
```
*[-ATR -high]
      |
*[-ATR +high]
      |
*[+ATR +low]
```

In this theory, the notion of cost is derived essentially from how deep of a filter must be switched off to permit the segment type in question. Low [+ATR] vowels are highly costly (i.e. highly complex) since the lowest filter must be switched off to permit them to occur.

A similar sort of hierarchy may now be proposed for the person features. We have observed that Mam requires reference to both [-I] and [-you], and Algonquian requires reference to [-you], but not to [-I]. Most languages require reference to neither. The hierarchy in (51) derives this result:

(51)
```
*[-you]
   |
 *[-I]
```

These filters operate at Morphology, deleting the values [-you] and [-I] in most languages, for example the Semitic languages discussed in chapter 1. This leaves these languages with only [+I] and [+you], which I abbreviated as '1' and '2' in chapter 1.

However, in some languages, at a certain cost, the filters in (51) may be switched off. Both filters are switched off in Mam, such that [-I] and [-you] are not impoverished at

Morphology. Only the lower filter is switched off in Algonquian, such that [-you] is still visible, while [-I] is deleted.

Because of the cost associated with switching off filters (a child will be prompted to do so only when positive evidence is provided), the filter theory derives automatically the result that Algonquian is an unusual (cross-linguistically 'marked') system, and the Mam system even more unusual. In the morphological realm, the Mam person system is unusual in the same way that a language with a [±ATR] distinction in low vowels is unusual.

Next, consider the filter which neutralizes 1st inclusive and exclusive: *[+I αyou]. Suppose that this filter is not as deep as the *[-you] filter, so that the three filters can be arranged as below:

(52)

*[+I αyou]
|
*[-you]
|
*[-I]

Observe now that, owing to the hierarchy of filters, any language which has no inclusive/exclusive distinction will also not have [-you] visible in Morphology. This follows because the *[-you] filter is deeper than the *[+I αyou] filter. Therefore the *[+I αyou] filter can be rewritten *[+I +you]. The value [-you] will be deleted by the lower filter, so the upper filter need refer only to [+you]. The hierarchy can now be written as in (53). Example languages are given at right. If the name appears lower than the filter, then the filter is switched off in the that language.

(53)

	English, Arabic, other Semitic
*[+I +you]	
|	Nunggubuyu, nearly all Australian
*[-you]	
|	Cree, other Algonquian
*[-I]	
	Mam, Sursurunga, other Austronesian

The English type has no inclusive/exclusive distinction so *[+I +you] is switched on. The Nunggubuyu type has an inclusive/exclusive distinction, but [-you] is not visible. Cree has in inclusive/exclusive and [-you] is visible, but not [-I]. Finally, Mam has an inclusive/exclusive distinction and both [-I] and [-you] are visible. The same pattern recurs in Sursurunga and other Austronesian, to be discussed in section 2.1.9

One consequence of this hierarchy is that reference to [-I] and [-you] is not only highly costly but also dependent on the existence of an inclusive/exclusive distinction in the language. I now consider two apparent counterexamples to this claim and show how these are in fact consistent with the hierarchy of filters.

2.1.8 Two Cases of Person Syncretism

Since Jakobson's (1942) article on the Paleosiberian languages, it has become better known that some of the languages of Siberia display unusual syncretisms in the person system, but lack the inclusive/exclusive distinction. In this section I focus on two of these, showing that neither [-I] nor [-you] need be invoked to express these syncretisms.

The first case comes from Chukchi, a Chukotko-Kamchatkan language described in Skorik (1977). The present and future terminative paradigms in the singular of a simple intransitive verb are provided below (capitals represent harmonizing vowels):

(54)	a. Present terminative		Example	Gloss
1	t(ə)-	-rkə-n	tə-kətgəntat-ə-rkə-n	'I run'
2		-rkə-n	kətgəntat-ə-rkə-n	'you(s.) run'
3		-rkə-n	kətgəntat-ə-rkə-n	'he runs'
b.	Future terminative			
1	t-rE-	-gʔE	t-ra-kətgəntat-gʔa	'I will run'
2	rE-	-gʔE	ra-kətgəntat-gʔa	'you (s.) run'
3	rE-	-gʔE	ra-kətgəntat-gʔa	'he will run'

As can be seen above, the first person category is marked with the prefix t(ə)-, whereas the 2nd and 3rd persons have no (overt) prefix and are homophonous. Although the [-I] categories, i.e. 2nd and 3rd person, are homophonous, the value [-I] need not be visible in Morphology since these categories form the elsewhere class that has no prefix.

Mere homophony of forms does not entail that the categories grouped together by homophony must be a natural class: they may be morphologically unmarked or the elsewhere class.

The second case comes from Ket, a Yeniseyan language spoken in central Siberia. Jakobson (1942:617) asserted that in Ket "the first and third persons tend to fuse." This seems to suggest that [-you] must be visible in this language. In fact, the 1st and 3rd categories end up being homophonous only in certain paradigms where truncation has neutralized them, according to the data in Krejnovič (1968c:23).

In paradigms of the so-called D-class, both 1st and 3rd person show the prefix d- as opposed to second person, which has k-. However, d- is limited to the masculine 3rd person. A comparison with the present tense paradigm shows that where vowel truncation has not taken place, the 1st and 3rd persons show di- and du- respectively:

(55)	a.	di-l'oqŋ	'I shake / я трясусь'
		ku-l'oqŋ	'you (s.) shake / ты трясешься'
		du-l'oqŋ	'he shakes / он трясется'
		dɤ-l'ɔqŋ	'she shakes / она трясется'
	b.	d-i-l'-l'oqŋ	'I shook / я тряссся'
		k-i-l'-l'oqŋ	'you (s.) shook / ты тряссся'
		d-i-l'-l'oqŋ	'he shook/ он тряссся'
		dæ-i-l'-l'oqŋ	'she shook / она тряссся'

The truncated prefix d-/k- occurs only when not immediately adjacent to the verb root (Krejnovič 1968c:53). In the past tense paradigm (55b), the interfix -i- and the tense morph -l'- 'past' intervene, and so the truncated prefixes occur.

Jakobson's 'fusion' is better understood as a morphophonological accident rather than a deeper neutralization. No reference to [-you] need be made.

These two examples from Chukchi and Ket reveal that mere homophony of forms is not sufficient grounds for postulating the homophonous forms as a natural class definable by a feature visible at Morphology. Homophonies may arise as unmarked forms, elsewhere classes, or through morphophonological processes that neutralize distinct affixes at the phonological level. Neither the Ket nor the Chukchi system of person marking poses a problem for the hierarchy of filters proposed in the previous section.

2.1.9 On the Markedness of [±participant]

If the values for [±I] and [±you] are known, the value for [±participant] is redundant:

(56)
a. [+I] --> [+participant]
b. [+you] --> [+participant]
c. [-I -you] <---> [-participant]

Of these redundancies, only (56c) is a two-way implication. Now given the redundancies in (56a,b), the value [+participant] can be understood as an *enhancing* value for [+I] or for [+you].

Recall from section 0.2.2.4 and section 1.14 that I argued that the value [-singular] is a redundant enhancement of [dual]. Sometimes a dual is morphologically enhanced by [-singular] and patterns with the plurals, otherwise, the dual is not morphologically enhanced by [-singular], in which case the dual patterns with the unmarked number, typically singular.

We have already seen evidence from Winnebago, Navajo, and Lummi that the [+participant] categories may function as a natural class in morphology. I now show that the 1st and 2nd person categories may be enhanced by a [+participant] marker, while remaining distinct in virtue of other markers.

The language Sursurunga (Hutchisson 1986), spoken in southern New Ireland, Papua New Guinea, displays not only a dual and trial numbers but also a quadral number, for referring to four persons. The quadral number has a variety of other uses which are described in Hutchisson's study of the use of quadrals. I will return to this in a later section. What is of interest now is that in Sursurunga, like many Austronesian languages and in some ways like Mam, the 1st person exclusive and the 2nd person pattern together morphologically. The 1st inclusive has a special mark, and, in the dual, patterns with the 3rd person:

(57) Sursurunga Pronominal Forms (Hutchisson 1986:5)

	singular			plural	dual	trial	quadral
	A	B	C				
[+I -you]	-ng	i	iau	gi-m	gi-Ø-ur	gi-m-tul	gi-m-at
[+I +you]	*	*	*	gi-t	gi-t-ar	gi-t-tul	gi-t-at
[-I +you]	-m	u	i?au	ga-m	ga-Ø-ur	ga-m-tul	ga-m-at
[-I -you]	-n	a	*	di-Ø	di-Ø-ar	di-Ø-tul	di-Ø-at

I will be concentrating only on the nonsingular categories. (The singular categories show three forms, A, B and C, whose distribution is roughly a matter of case). For the first element of the nonsingular forms, the following rules suffice:

(58) a. -I -you -sg di-
b. +I -sg gi-
c. +you -sg ga-

Recall that because [±I] is higher than [±you] on the hierarchy of features of section 2.1.4, then (58b) will be ordered before (58c). This gives the result that [+I +you] will appear with prefix gi- by (58b), which applies first.

Turning now to the second morph, observe that only the 1st inclusive [+I +you] category has *-t-*. The remaining participant categories, i.e. 1st exclusive and 2nd person, have *-m-* in the plural, trial and quadral, and -Ø- in the dual. Hutchisson analyzes -t- are the mark of the inclusive and *-m-* (or -Ø-) as the mark of the exclusive, by which he means "excludes either hearer or speaker, but not both; only the person category indicated [by the prefix] (p.4)". In other words, the "exclusive" is either [-I] or [-you], but not both. Finally, the [-participant] category has neither *-m-* or *-t-* but realizes this second position as -Ø-.

By analogy to the analysis for Mam, one might assume that the "exclusive" of Sursurunga includes those categories where the person feature values disagree.[7]

However, this is not necessary here, since the exclusive categories are all and only those [+participant] categories remaining after [+I +you] has its own mark, namely *-t-*. The exclusive categories receive the affix *-m-* (or -Ø- in the dual) by (59f,g) below.

(59)
a. +you (+I) -t-
b. +participant dual -Ø-
c. +participant -m-
d. Elsewhere -Ø-

Rules (59b,c) refer to the value [+participant], which enhances [+I] or [+you]. Thus, in such forms as *gi-m* 'we excl.' or *ga-m* 'you (pl.)', the Stem Extension *-m-* marks these

[7]Nogogu, a Melanesian language (Forchheimer 1953:81, cf. Zwicky 1977) has a pattern of pronouns reminiscent of Sursurunga:

	singular	dual	trial	plural
[+I -you]	(i) nou	om-o-rua	om-o-tolu	em-am
[+I +you]	*	o-rua	o-tolu	rie
[-I +you]	i niko	om-rua	om-tolu	em-iu
[-I -you]	i nikin	ru-rua	ritolu	i rir/rire (cont.)

The 1st exclusive and 2nd person are characterized by the prefix om- in the dual and trial and by em- in the plural. On the other hand, the 1st inclusive classes with the 1st exclusive, since both show o-. The distribution of om- is the exclusive categories in Hutchisson's sense, i.e. those categories where the values from [±I] and [±you] disagree.

Hutchisson notes (fn.4, p.20) that there are "at least four other New Ireland languages (Tigak, Madak, Kara, Tungak) which have one morpheme for 1+2, another for 1+1 and 2."

forms (redundantly) as [+participant] forms. This situation is analogous to the redundant enhancement of dual by [-singular] in Arabic *katab-tum-aa* 'write-2pl-dual.'

Finally the various cardinality marks fill the third position of autonomous word structure. Hutchisson notes (1986:20 fn.5) that *-ar* is the affixal form of the number *ru* 'two.' The suffixes *-tul* and *-at* are homophonous with the numbers 'three' and 'four.' Thus, it makes sense to assume that these suffixes are actually independent syntactic heads which incorporate into the pronominal (Determiner):

(60) $[_D\ D\ [_{Num}\ \text{Number}]]$

Stem $Q_{Extension}$ (Q_{Number})

After this head-movement, the X^0 of Number will by default license another position of exponence in the pronominal word-structure. This gives the third position to fill in the dual, trial, and quadral forms, but no additional position in the singular and plural. This final position is realized by the following affixes:

(61)
a. two (αI $-\alpha$you) -ur
b. two -ar
c. three -tul
d. four -at

Note that the distribution of the allomorphs of 'two' depends on the (previously discharged) features of [±I] and [±you]. The vowel raises to *-u-* only where the person features disagree, i.e. the 1st exclusive and second person. Otherwise, 'two' appears in its regular affixal from *-ar*.

Despite such cases where redundant [+participant] enhances [+I] or [+you], the more common situation is for the value [+participant] to be otiose in Morphology. Similarly, the value [-participant], which defines the 3rd person, is often otiose as well, giving the typical circumstance in which the 3rd person category is morphologically unmarked. I will have little to say about this except that it appears that filters of the sort *[+participant] and *[-participant] are regularly assumed. Rules referring to either value of [±participant] are therefore somewhat costly, although exactly how much remains an open question.

2.1.10 Summary

In the preceding section I have proposed a theory of person features [±I], [±you] and [±participant]. These features, along with a simple feature hierarchy (I > you) suffice to permit all and only those person categories observed for natural language. Certain more complex categories can be built up by incorporating one pronoun into another, as will be discussed after the number features are introduced.

I presented evidence from Algonquian, Mam and Sursurunga that the values [-I] and [-you] must be morphologically visible at least in some languages. However, I proposed that reference to these values, and possibly to [±participant], is more costly than reference to [+I] and [+you], which are universally available morphologically. To encode cost in the theory, I proposed a hierarchy of filters like those found in Calabrese's theory of phonological inventories. According to this filter, reference to [-I] is most costly and implies (potential) reference to [-you] in the system. Reference to [-you] is a bit less costly and implies the presence of an inclusive/exclusive distinction. Languages without an inclusive/exclusive distinction are predicted to never have [-I] or [-you] active morphologically. Apparent counterexamples from Chukchi ([-I]) and Ket ([-you]) were argued to be elsewhere cases or mere phonological homophonies.

Certain difficulties remain with the theory, however. Observe that in the languages which require both [-I] and [-you] to be visible (Mam and Sursurunga, for example), the rules are in both cases α-rules referring to both values of [±I] and [±you] at once. Whether the values for person features agree or not is apparently a morphologically relevant piece of information. Similarly, in the Algonquian case, reference to [-you] needs to be invoked only in a rule which also refers to [+I]. In no case are the classes [-I] or [-you] picked out as *independent* classes. This suggests that the theory of markedness I have proposed does not fully articulate the cost of morphological representations. I raise this complex issue here, leaving it open for further work.

The need for α-notation rules to express the distribution of the enclitics in Mam is of general theoretical interest in establishing this formal mechanism within Morphology. In sections 2.3 I will show that α-notation is also necessary for the number features to account for the Number-Switch rule of Kiowa-Tanoan.

2.2 Number Features

The following several sections treat the number features. At first glance, the number features would seem to be relatively straightforward indices of the cardinality of an argument. However, there are certain complexities beyond this. The most significant is the interaction of what I will call the *inherent* number of an argument and its explicit, or morphologically marked number. In section 2.2.1 I show that 1st person inclusives, being inherently dual, interact with explicit number in interesting ways. Following Conklin (1962), I propose that there is a number feature, [±augmented], which is interpreted through the (proper) subset relation. Evidence is presented in 2.2.2 that if both [±singular] and [±augmented] are active in Hopi, the distribution of suppletion and agreement allomorphy and reduplication is completely uncomplicated. In section 2.2.3 I examine composite pronouns (pronouns formed by syntactic incorporation). I show that the stem of a composite pronoun is always [+augmented], and that even the most baroque systems of composite pronouns, found in the Bantu language Bamileke, are derived by combination of simpler categories. The interpretation of complex pronouns combines both the explicit, truth-functional semantics of a form with functional inferences made by speakers as to the felicitous use of the form. The interaction of inherent number and functional inference is then illustrated with the pronominal system of Nunggubuyu in section 2.2.6. Then, in section 2.2.7, I show that inherent number and functional inference explain the so-called 1st person Limited Inclusive in Sierra Popoluca, which is a prima facie counterexample to the claim that there is no feature [-3]. Finally, in sections 2.3 - 2.3.3, I argue that both inherent number of nouns and a syntactically-supplied class feature [INVERSE] interact in the Kiowa-Tanoan languages to give a complex relationship between semantic, morphological, and classificatory (agreement-relevant) number-class.

2.2.0 The Simple Cardinality Features

In many languages the number features are uncomplicated and are mapped to simple rules governing the cardinality of the index set of the relevant argument:

(62)
a. [+sg] ---> 1
b. [-sg] ---> > 1
c. dual --> 2
d. trial --> 3
e. quadral --> 4

Beyond these relatively trivial observations, the number features do have interesting properties. First, while both values of [±sg] are morphologically active, defining the traditional classes *singular* and *plural*, all the evidence I have encountered is consistent with the hypothesis that the cardinality features [dual], [trial] and [quadral] are monovalent. In other words, the complements of these classes, e.g. [-trial], do not function as natural classes.

In terms of markedness, the value [+sg] is more costly than the value [-sg]. We can then hypothesize that any language which has [+sg] morphologically active will also permit reference to [-sg], but a less costly system will impoverish [+sg] by means of a filter *[+sg]. This accords with the usual expectation that singulars are an unmarked category vis-à-vis plurals.

Second, it has been well known since Greenberg's (1963) study of Universals of Human Language that if a language has a trial number it must have a dual (Universal 34, p. 112); in the case of the Sursurunga quadral, it appears that quadral implies trial and dual. The following hierarchy of filters expresses this:

(63)
*[dual]
|
*[trial]
|
*[quadral]

However, what has been labeled as trial, dual and plural are not always of cardinality 3, 2, and more than 1 respectively. In the next several sections I will examine a number of number systems with an eye to determining their true morphological nature.

2.2.1 Number and the 1st Person Inclusive

Because [+I] and [+you] each guarantee that one individual must be involved, the 1st person inclusive is *dual by necessity*. We may then refer to the 1st person inclusive as *inherently* dual, and, by implication, inherently nonsingular. From a morphological point of view, however, 1st inclusives sometimes pattern with what could be termed the singular categories while at other times with the plurals, i.e. nonsingulars. This leads to what Greenberg (1988) has called the "ambiguous" nature of the 1st inclusive.

Not every language which has a 1st person inclusive has a specifically dual number. In other words, although 1st inclusive [+I +you] is inherently dual, having a this category does not imply that [dual] is an active feature of the system. There is no implicational relation between the filters *[+I +you] and *[dual]. Arabic or Meħri, for example, have an active dual without an inclusive. In these languages the feature [dual] is not impoverished, and the dual category has an explicit mark, e.g. *-aa(ni)* in the Arabic prefix-conjugation. We may call these *explicit* duals. The Algonquian languages, on the other hand, have *[dual], while permitting [+I +you]. Dual in these languages is strictly inherent.

Consider first an instance in which inherent [+I +you] patterns with explicit dual morphologically. The following table shows the pronominal system of Worora, a non-Pama-Nyungan language of Australia (data from Forchheimer 1953:126):

(64) Pronouns in Worora

	sing.	dual	trial	plural
[+I]	ŋaiu	ar-endu	ari-ŋguri	ari
[+I +you]	*	ŋar-endu	ŋari-ŋguri	ŋari
[+you]	ŋundju	njir-endu	njiri-ŋguri	njiri

Observe that the 1st person inclusive category behaves just like the 1st exclusive and 2nd person. Although inherently dual, the [+I +you] stem ŋar(i)- may be explicitly marked as dual by the suffix -endu. Importantly, the [+I +you] forms do not pattern with the singulars. In the array above, the absence of a "singular" 1st inclusive is shown by the gap with the *.

On the other hand, the subject pronoun affixes of Winnebago (Lipkind 1945, Ken Hale, p.c.) present an instance where the 1st inclusive patterns with the singular arguments of the other persons:

(65) Winnebago Person-Number Affixes

	nonplural	plural		category name
[+I +you]	hin-	hin-	-wi	1st inclusive
[+I]	ha-	ha-	-wi	1st exclusive
[+you]	ra-	ra-	-wi	2nd
[-prt]	ø		-ire	3rd

Patterns such as this are not in fact uncommon, although their occurrence is limited to languages which are less studied.[8] Conklin (1962) first remarked on this sort of pattern in Hanunoo (Philippines), and proposed a feature [±restricted]. Similarly Matthews (1972a) observed the same pattern in Huave (although it is morphologically obscured), and postulated feature [±others].[9] The [+restricted] or [-others] arguments are in the left column above, the [-restricted] or [+others] in the right column. The interpretation of this feature is: the [+restricted] forms have only the inherent number of the argument, while the [-restricted] forms have at least one more than the inherent number.

Consider the morphologically "plural" or [-restricted] 1st inclusive ha- -wi in Winnebago. Because this form is inherently dual, the [-restricted] form must have at least one more than dual, i.e. at least 3. From the point of view of strict cardinality, the "singular" 1st inclusive has 2 while the plural 1st inclusive has at least 3; all other arguments have the normal singular = 1 and plural = >1. To ensure that semantic interpretation proceeds properly, the [+restricted] forms cannot be [+sg], for this would demand a cardinality of 1 for [+I +you]; similarly the [-restricted] forms cannot be [-sg] as interpreted as "greater than 1," since this mistakenly would permit the [-restricted] 1st inclusive to mean as few as 2, when it actually means at least 3.

Because of such problems, I will adopt the feature [±augmented], following the usage of McKay's (1981) study of Rembarrnga, a non-Pama-Nyungan language of No. Territory, Australia. [±augmented] is the inverse of Conklin's [±restricted]: the left column in in (65) will be [-augmented], while the right column will be [+augmented]. The choice of this terminology will become clear when I discuss composite pronouns in section 2.2.3.

I propose a formal interpretation of augmentation as follows. First, it is clear that the additional index implicated by augmentation cannot be paired with a discourse role forbidden by some other feature of the augmented argument. For example, the additional individual-concept in a 2nd person [-I +you +aug] argument cannot be the speaker owing to [-I]. Hence, augmentation is dependent in its interpretation on the values of the other features. We express this as follows:

---

[8]Greenberg (1988:2) lists languages with 8-person systems like that of Winnebago. Among the better known ones are Southern Paiute, Miwok, Ilocano, and Margi. But having an 8-person system is not necessarily the same as showing the pattern of morphological distribution shown in Winnebago, with a particular affix associated with the [-restricted] class. For example, Margi has an 8-person system but does not have an explicit augmented affix.

[9]In accordance with the idea that the semantic interpretation of person-number features provides only the minimal conditions under which a pronoun my be used, it would seen better to call Matthews' feature [±other] rather than [±others]. This is because a [+other] category requires at least one other individual, although more than one is of course also possible depending on any further cardinality restrictions such as [trial].

(66) a. *A*, [+aug]: If N is the set of anaphoric indices borne by *A*, then N must have some nonempty proper subset which fulfills the conditions imposed by the other features of *A*.

b. *A*, [-aug]: If N is the set of anaphoric indices borne by *A*, then no nonempty proper subset of N fulfills the conditions imposed by the other features of *A*.

Inversely, [-augmented] categories are such that *no* proper subset of N fulfills the conditions that N must fulfill. Hence, a [-augmented] argument must have at most the minimal assignment sufficient for fulfilling the conditions imposed by its other features. As such, [-augmented] categories are "restricted" in Conklin's original sense, or "minimal" in the sense of McKay (1981), among others.

Including the feature [±augmented] in the inventory of universal features of number is a significant expansion of the theory of number features. One is inclined to consider whether both [±sg] and [±aug] are necessary in Universal Grammar. I will first show that neither can be reduced to the other, and, more compelling still, that both can be active simultaneously.

2.2.2 Augmentation vs. Nonsingularity

Suppose first that there is in fact no [±sg], but in fact all singular:plural oppositions are encoded by means of the proper subset relation of [±augmented]. This has no immediately bad consequences for language without 1st inclusives. For example, English *we* and *they* can be interpreted either as [-sg] or as [+aug]. However, returning to the example from Worora (64), a language with 1st inclusives, observe that all the [-sg] arguments form a class of stems ending in -r(i). Suppose that [-sg] were replaced by [+aug] in all instances. The class of pronouns with stems in -r(i) can no longer be defined, since ŋar- would be [+I +you -aug] in the dual but [+I +you +aug] in the trial and plural. This pattern shows that [±sg] is active in Worora and [±aug] is not.

On the other hand, the fact that 1st inclusive sometimes patterns with the [+sg] arguments, as in Winnebago, implies that [-aug] cannot be reduced to [+sg].

If both [±sg] and [±aug] are available, then it is predicted that some language should make use of both. This prediction is borne out by the distribution of suffixal morphology, suppletion and reduplication in Hopi.

Ken Hale (to appear) has shown that the so-called dual category in Hopi forms the intersection between the singular and plural categories, as illustrated in the following sentences:

(67) a. Pam wari
that run-PERF 'He/she ran.'

b. Puma wari
those run-PERF 'They (two) ran.'

c. Puma yu?tu
those run-PERF 'They (plural) ran.'

(68) a. Pam taaqa wari
that man run-PERF 'That man ran.'

b. Puma taaqa-t wari
those man run-PERF 'Those two men ran.'

c. Puma ta?taq-t yu?tu
those men run-PERF 'Those men (plural) ran.'

First, examine (67). There is a two-way opposition in subject pronouns and a two-way opposition in verbs. The verb meaning 'run' shows suppletion for number: the suppletive 'plural' form occurs only in (67c), but the nonsingular subject pronominal occurs in (67b) and (67c). When the nonsingular pronominal cooccurs with the nonplural verb in (67b), a dual interpretation results.

In (68b) the situation is similar. The nominal argument 'man' shows number two ways. The suffix -t has the same distribution as the nonsingular form puma; but 'plural' reduplication (ta?taq-) occurs only in (68c), correlating with the suppletive plural verb form. Where -t occurs without reduplication, the interpretation is dual.

This situation is easily accomodated within a system with both [±sing] and [±aug]:

(69) Hopi Number Categories

label	[±sing]	[±aug]
singular	+	-
dual	-	-
plural	-	+

The dual number in Hopi is really the "minimal" nonsingular category. In the terms presented here, minimal translates to [-aug]. We are now in a position to explain precisely what the plural category is: a plural argument has a nonsingleton index set, (hence minimally 2) augmented by at least 1 (hence minimally 3). The morphological facts follow straightforwardly. Nominal reduplication and verbal suppletion are conditioned by [+aug]; the nominal suffix -t and the nonsingular pronominal puma 'those,' reflect [-sing].

The status of dual as [-sg -aug] is confirmed in the Kiowa-Tanoan languages as well. These more complicated systems will be treated extensively in separate sections (2.3-2.3.3).

2.2.3 Composite Pronouns

The pronominal system of Rembarrnga illustrates the phenomenon of composite pronouns. Composite pronouns are essentially compounds where one member is an augmented pronoun stem and the second member is an incorporated pronoun which delimits the augmentation. Consider the following set of dative pronouns in Rembarrnga from McKay (1981):

(70) Rembarrgna Dative Pronouns

	minimal	unit augmented	augmented
1	ŋənə	yarr-pparraʔ	yarrə
1+2	yəkkə	ŋakorr-parraʔ	ŋakorrə
2	kə	nakorr-parraʔ	nakorrə
3 m	nawə	parr-pparraʔ	parrə
3 f	ŋatə	id.	id.

The 'unit augmented' forms are augmented by only one. Hence, the 1st exclusive unit augmented refers to 2 individuals ('to me and him/her'), but the 1st inclusive unit augmented refers to 3 ('to me and them'). The regular augmented forms on the far right are interpreted as one more than the unit augmented forms.

In terms of features, the regular augmented forms are simply [+aug]. Their interpretation as *one more* than the unit augmented forms must come by way of functional inference. That is, the hearer will assume that a regular augmented form refers to one more than the unit augmented forms because if it did not, the speaker would have indeed used the unit augmented form. The regular augmented forms are probably also used in situations where vagueness is desired or precision unnecessary, just as a speaker of English might loosely use *all* instead of *both* to refer to 2 individuals or things. I will return to the role of functional inference later.

Observe that the unit augmented forms are all derived via suffixation of -(p)parraʔ to the regular augmented forms in the right column (minus final schwa). The suffix -(p)parraʔ delimits the augmentation of the category to just one.

For example, the unit augmented 1st exclusive has the augmented stem yarr- [+I +aug], where the augmentation is delimited to one individual. The stem requires at least {1} + {x}, and the suffix ensures that {x} has cardinality 1.

The unit augmented forms are not, strictly speaking, explicit markers of *cardinality*, but rather explicit delimitations of the augmentation of the stem.

All languages which display composite pronouns show stems which are augmented in form, followed by some pronoun showing the delimitation of this augmentation. Such composite forms are typical of many Bantu languages and also occur in Ewe (Forchheimer 1953:132). Hyman (1979:47-55) provides a good discussion of composite pronouns in Aghem, a Western Grassfields Bantu language of the Ring group (Northwest Cameroon).

An extremely baroque system of composite pronouns occurs in Bamileke, a Grassfields Bantu language of the Mbam-Nkam group spoken in Cameroon northwest of Douala. Wiesemann (1986:viii) provides an extremely rich inventory of Bamileke pronouns for the Ɣɔmáláʔ dialect of Bamileke. The following chart, to be explained momentarily, displays the full set of forms:

(71) Bamileke-Ɣɔmáláʔ Pronouns

	-aug	+aug	3∈N	N ∪ 3	N ∪ 3 +aug	N ∪ 2
[+I -you]	gə	pyə	pyə-ė	py-ayʉ̇	pyə-apu	py-awu
[+I +you]	pu	pə	---	pə-ayʉ̇	pə-apu	---
[-I +you]	o	po	po-ė	po-ayʉ̇	po-apu	---
[-I -you]	e	wap	pu-ė	---	pu-apu	---

The various columns have the following interpretations. The [-aug] and [+aug] columns have their regular interpretations as described in the last section. Observe that all the composite pronouns in the remaining columns have the [+aug] forms as stems, except the 3rd person [-I -you], which has a special composite stem pu, which happens to be homophonous with the unaugmented [+I +you] pu.

The forms in the 3 ∈ N column are augmented forms where the augmentation is explicitly a single third person. Thus pyə-ė means 'we (excl.) inc[illegible]g him/her.' In other words, 'him/her' is *contained* in 'we.' Such a phrase would be analogous to the Russian *my s nim/nej* мы с ним/ней, meaning not 'we and he/she,' but rather 'we, among whom is he/she.'

The forms in the $N \cup 3$ column are the augmented forms plus a single third person. Thus py-ayʉ means 'we (excl.) and him/her.' In other words, 'him/her' is added to the set comprising 'we.' The Russian analogue would be *my i on/ona* мы и он/она, 'we and he/she.' The difference between pyə-é and py-ayʉ is the difference between set-inclusion and set-union.

The forms in the fourth column appear to have readings in which an augmented 3rd person set is added to the augmented stem set. Thus pyə-apu would appear to mean 'we (excl.) and them,' although it is unclear whether 'them' may be included in 'we' as well.

Finally the form in the last column has the addressee added to a 1st exclusive set.

The forms which are incorporated into the augmented stem have independent functions in the language. For example, -é in po-é 'you (aug.) including him' is clearly related to the 3rd unaugmented pronoun e 'he/she.' The element -ayʉ is clearly related to the emphatic pronoun yʉ 'he himself (lui-même)' (Ntagne & Sop 1972:37). The suffix -apu comes from the augmented stem pu. (The form puapu is not reported in Wiesemann 1986 but appears in Ntagne & Sop 1972:38). Finally the suffix -awu is related to a variant of the 2nd person singular pronoun, namely wù (Ntagne & Sop 1972:41).

There is some unclarity in reporting the meanings of these composite pronouns. In the table below I compare the glosses from Wiesemann (1986) and Ntagne & Sop (1972):

(72)	form	gloss (Wiesemann)	gloss (Ntagne & Sop)
	pyə-é	'I and one absent'	'lui et moi'
	po-é	'thou and one absent'	'vous (2)'
	pu-é	two absent persons representing 2 parties	'ils (2)'
	py-ayʉ	'we and him'	not reported
	pə-ayʉ	'I, thou, one absent' or 'we, thou, one absent' or 'we, you all and one absent'	not reported
	po-ayʉ	'you all and him'	not reported
	pyə-apu	'we and them'	'nous +eux -vous'
	pə-apu	'we, you all, and several absent persons'	'nous +eux +vous'
	po-apu	'you all and them'	'vous +eux'
	pu-apu	not reported	'ils +eux'
	py-awu	'we and thou'	not reported

What remains unclear regarding these glosses is whether, for example, in a combination such as pə-ayʉ the 'thou' or 'you all' of the gloss may contain individuals which are also simultaneously part of the 'we' of the gloss. For example, it makes sense to suppose that pə-ayʉ has only one interpretation, namely *minimally* 'I, thou, and one absent,' but the various glosses reflect the fact that, in the non-minimal interpretation, all additional parties which may be involved except the emphasized 'one absent' may be part of the 'we' group and the 'you all' group.

Support for this conjecture comes from Voorhoeve's (1967) discussion of complex pronouns in another dialect of Bamileke, as spoken at Bangangté. It is clear from Voorhoeve's glosses that the pronouns system of this dialect is as below:[10]

(73) Bamileke Pronouns (Voorhoeve 1967)

	-aug	+aug	$3 \in N$	3 aug C N	$2 \in N$	2 aug C N
[+I -you]	mə	bag	bâg-jé	băg-à-bo	bág-uɨ	bág-à-bìn´
[-I +you]	o	bin	bîn-jé	bín-à-bo		
[-I -you]	a	bo	bô-jé	bó-à-bo		

Unlike ɣɔmálá?, this dialect of Bamileke nas no 1st inclusive form. For ɣɔmálá? pə 'we (incl.)', this dialect of Bamileke uses a composite pronoun bâg-uɨ, which incorporates a 2 sg pronoun into the stem of 1st exclusive form. This leads to a certain interpretive paradox: namely $2 \in \{1,3\}$. However, since this dialect of Bamileke has no specifically inclusive 1st person forms, this paradox may be resolved if the exclusive form is merely [+I Øyou] when functioning as a stem in composite pronouns.

Consider the glosses which Voorhoeve provides from the composite pronouns, and an interpretation of this gloss in set notation at right:

[10]Voorhoeve also gives (p. 427) four reciprocal pronouns: băg-bag 'we between ourselves excl.,' bə̆n-bən 'we between ourselves incl.,' bĭn-bìn 'you pl. between yourselves,' and bó-bó 'they between themselves.'

(74)	bâg-jé	'I plus him' (dual)	{1,3} dual
	bîn-jé	'you (sg.) plus him' (dual)	{2,3} dual
	bô-jé	'he plus him' (dual)	{3} dual
	bâg-uɨ	'I plus you sg.' (dual)	{1, 2} dual
	băg-à-bo	'I plus them or we plus him/them'	{1, 3} 3 or more
	bĭn-à-bo	'you sg. plus them or you pl. plus him/them'	{2, 3} 3 or more
	bó-à-bo	'he plus them or they plus him/them'	{3} 3 or more
	bág-à-bĭn	'I plus you pl. or we plus you sg./pl.'	{1, 2 (3)} 3 or more

What is important here is the glosses of the last five in the above list. The various English glosses are merely different ways of expressing the conditions of the set notation at right. For example, bĭn-à-bo means 'you sg. plus them,' i.e. (2) + (3,3 ...) or 'you pl. plus him,' i.e. (2, x...) + (3), where x = 2 or 3, or 'you pl. plus them' (2, x) + (3, 3 ...), where x = 2 or 3. All the English glosses are consistent with the condition that both the roles {2} and {3} must be represented *and* there must be at least three participants. This pattern is predicted if the incorporated pronoun is interpreted as a subset of the pronoun into which it is incorporated, as reflected in the notation in the chart in (73).
A more thorough investigation of the truth-conditions on the use of the complex pronouns in Ɣɔmálá? would probably reveal a similar pattern as in Voorhoeve's study.

Because of the additive and delimiting nature of the the incorporating pronouns, it is possible in Bamileke to denote very specific groups of individuals if necessary. For example, the Ɣɔmálá? form po-é, meaning 'thou and him/her' is a form {2,3} with specific inclusion of a third person.[11] The pronoun po-é is not used to denote two addressees: for this a simple po 'you (pl.)' would be used. Such very specific forms like po-é appear at first glance to flout Generalization II (section 2.1.1), i.e. that {3} is always an optional role except in the 3rd person. In fact, however, it is no accident that such very specific forms occur only as composite pronouns where two independent categories accumulate their discourse roles. Only by such an accumulation of simpler categories could such a form arise. A specifically {2,3} form could never be monomorphemic.

It should be borne in mind that a form such as po 'you (pl.)' probably does not truth-functionally exclude an interpretation of {2,3}. This follows from Generalization I: no category specifically excludes {3}. Yet if a speaker wishes to explicitly denote the class

[11]Presumably such forms prompted Greenberg to use the phrase "rare and doubtful exceptions" when stating the universal regarding 2nd person exclusives.

{2,3}, then the form po-é could be used. We may then conclude that po alone is either vague, or functionally inferred to be {2} without {3}.

Although the precise interpretation and use of composite pronouns in Bamileke is far from well understood, we may say in summary that the system described by Voorhoeve and the ɣɔmálá? system both greatly exploit the possibilities of incorporating pronouns into others to form composite entities with complex semantic interpretations. However, even such richly expanded inventories are built up from the simple categories defined by the person and number features proposed here.

2.2.4 Functional Inference and Ambiguity

Greenberg's discovery of the terminological ambiguity surrounding the 1st person dual amounts, when properly understood, to an acknowledgment that person-number *categories* are not primitives, but rather the features which define the categories are the primitives. Consider Greenberg's discussion:

> There is here a kind of paradox. I first sought to show that a relativistic view in which the conceptual system implied by a three-feature analysis is assumed to be different and perhaps incommensurate with that implied by the 'traditional' one is brought into question by diachronic evidence. However in doing this we have sacrificed the notion of a *uniform and universally valid set of typological categories* by positing an ambiguous one (Greenberg 1988:12, italics mine).

Abandoning the ontological priority of categories in favor of the features that compose them, as here, renders this a pseudo-problem. Each speaker of a language constructs a system of categories composed from the features afforded by Universal Grammar combined with specific choices motivated by the learning stimulus. These choices are encoded by means of filters which dictate which features and combinations of features may be active in the system.

A speaker's knowledge of the system *in toto* permits further inferences of a functional nature. For example, in a language with a dual and a plural [-sg], the plurals are regularly interpreted as one more than the duals, even though strictly speaking, plurals are merely nonsingular, i.e. more than one. Such functional inferences can become highly complex, depending on the complexity of the system. In Bamileke, for example, it would appear that the extensive enlargement of the pronominal system by means of incorporation leads to rather fluid conditions on the appropriate use of the various less and more specific pronominal forms.

The existence of such functional or discourse-bound knowledge of pronominal use alongside the truth-functional interpretation of the forms leads to the plausible but incorrect view that the systemic function or position of a form is the same as its grammatical representation. This leads directly to Greenberg's paradox. An explicit refusal to conflate these types of knowledge characterizes generative phonology as opposed to structuralist phonology, and must underlie any explanatory morphosemantics of person and number as well.

Diachronic shifts from inherent to explicit duals are to be expected since the 1st person inclusive, being *functionally* ambiguous, can be acquired in more than one way, i.e. have more than one possible mental representation. Under the appropriate learning conditions, one representation can easily replace another.

2.2.5 Trials and Paucals

Because incorporation forming a composite pronouns may explicitly delimit cardinality, it may happen that a form will be functionally or inherently trial while not being explicitly trial. For example, the unit augmented 1st inclusive in Rembarrnga, ŋakorr-parraʔ 'to me and you (dual),' refers to three persons, but morphologically it has no explicit mark for the trial. There is no reason to suppose that a feature [trial] is active in Rembarrnga.

On the other hand, explicitly trial arguments, as far as I am aware, always show the incorporation of a numeral 'three' as in Sursurunga, *gi-m-tul*, where *-tul* is the number three. This same pattern recurs in Fijian (Anderson & Keenan 1985:263) and appears to be an archaic one in proto-Polynesian.

I have not found any instances in which the feature [trial], or for that matter [quadral], figures together with another feature in a rule of morphology. In other words, whenever [trial] has a specific mark, this mark is individually segmentable and never fuses with person and number. The same cannot be said of explicit duals, as is well attested in the fusion of dual number and person in the older Indo-European languages, e.g. Gothic, Sanskrit, Homeric Greek.

Because [trial] and [quadral] when explicitly marked are detachable in this way, it seems reasonable to suppose that trials and quadrals are formed syntactically through incorporation of a number X^0 within the DP headed by the pronominal stem. I illustrated this process in the discussion of Sursurunga.

Leaving aside then the cases of implicit trials (via composite pronoun incorporation) or explicit trials (via number incorporation), the one case of trial left is that of the paucal-trial.

Hercus (1966) shows that in the (now extinct) Australian languages of Victoria, the trial suffix in many cases derived historically from a word meaning 'a group of people standing or sitting together or associated with each other in some way.' (Hercus 1966:336). For example in Wembawemba, the suffixation of *-guli* 'crowd,' renders a pronoun trial:

(75) njagidj-min nja jandin-guli
what-for indeed us (pl. incl.)-crowd

'Why all three of us?" (Hercus 1966:336)

A similar situation arises in Arabana, once spoken near Lake Eyre, South Australia:

(76) a. gari 'group of people'
b. aniri 'we' (pl. i.e. more than three)
c. aniri-gari 'we three'
d. urgari 'you' (pl., i.e. more than three)
e. urgari-gari 'you three'
f. maŋgara-gari 'a group girls' (i.e. the constellation of the seven sisters)
g. gari-guna wadlu 'group-POSS country' =
'the (other) group of people's country, another tribal territory'

When the word meaning 'group of people' (76a) is suffixed to the plural pronouns in (76b,d), the corresponding trials result. (76f) attests that this same morph, when suffixed to an NP, means merely 'a group of ...' without a commitment to cardinality.

So explicit trial markers are often really 'paucal' markers, expressing various notions such as 'a few,' or 'a countable group.' For example, in Ngarinjin, a language closely related to Worora, the explicit trial suffix is *-njina* or *-na* and means 'three or a small number' (Coate & Oates 1970:28):

(77) Cardinal Pronouns in Ngarinjin

	[+sg]	[-sg, dual]	[-sg, trial]	[-sg]
[+I +you]	--	ŋad-njiri	ŋad-njina	ŋar-un
[+I]	ŋ-en, ŋ-in	njad-njiri	njad-njina	njar-un
[+you]	njaŋ-an, njiŋ-an	nud-njiri	nud-njina	nur-un
[-prt, gender i]	andu	b-andu-njiri	b-andu-njina	b-andu(n)
[-prt, gender ii]	nj-andu			
[-prt, gender iii]	w-andu			
[-prt, gender iv]	m-andu			

The nonsingular forms share a particular stem type (*-d* alternates with *-r*) in the 1st and 2nd person forms. The third nonsingular forms act like a gender in themselves (as in many Australian languages), such that only one gender is distinguished in the nonsingular, marked by the prefix *b-*. The morphologically plain nonsingular arguments in the extreme right column, called 'plurals,' manifest the [-sg] stem along with the [-sg] pronominalizing suffix *-un*. Whenever dual or 'trial' = paucal are explicitly marked, *-un* does not appear.

To summarize, the category trial, like the category dual, is ambiguous in Greenberg's sense. In some cases, trials arise inherently, owing to unit augmentation of the 1st inclusive form, as in Rembarrnga. In other cases, trials are paucals in disguise, as in Ngarinjin and the other Australian languages (Wembawemba, Arabana) cited by Hercus. Whenever trials are not inherent in one of these ways, they are explicitly marked by the number three incorporated into a [-sg] or [+aug] stem, as in Sursurunga, Fijian, and other Polynesian languages.

2.2.6 Inherent Dual and Trial in Nunggubuyu

In this section I present an analysis of the person-number categories of Nunggubuyu based on the excellent description of Heath (1984), to which I am indebted for the morphophonemic analysis. What follows in part supersedes an earlier account of mine presented in Noyer (1991). The discussion that follows also serves an introduction to the multiple argument agreement system of Nunggubuyu to be given in chapter 3. Nunggubuyu is a non-Pama-Nyungan language of Australia, spoken in the Northern Territory.

As in Rembarrnga, probably a related language, the Nunggubuyu pronominal system employs the [±augmented] distinction. As in Ngarinjin, plurals are marked by a special gender marker as well. As a result of the cooccurrence of these two types of number marking, 1st inclusive trials and 1st exclusive and 2nd duals end up being morphologically *unmarked* categories as I will show.

Recall that in Ngarinjin (77), 3rd person plurals show a particular plural gender marker *b-*. Nunggubuyu has extended this sort of system to the [+participant] categories as well. The plural gender marker is *-ru*. The plural gender in the 3rd person may be used to denote collectives or what might be glossed as singular inanimate nouns, as well as plural groups of humans. Thus the gender marked by *-ru* is not to be correlated with the number feature [-sg], but rather to a gender feature [PLUR] whose uses are not restricted to cardinality.

Consider the 1st person categories arrayed below (The coefficient 'm' in the chart means 'mark' and notates the presence of a privative feature. Little hinges on whether these features are regarded as privative or bivalent; see discussion in chapter 3):

(78) Nunggubuyu 1st person categories

category	intr. prefix	pronoun (nominative)	[± I]	[± you]	[± aug]	PLUR	MASC	FEM
1 sg	ŋa-	ŋaya	+	-	-			
1 excl m dual	nV: - n̠i-	nV-n̠i	+	-	+		m	
1 excl f dual	nV: - ŋi-	nV-rŋi	+	-	+			m
1 excl plur	nV: - ru-	nV-ru	+	-	+	m		
1 incl dual	na-	nagawa:	+	+	-			
1 incl m trial	ŋV: - ni-	ŋa-gu-n̠i	+	+	+		m	
1 incl f trial	ŋV: - ŋi-	ŋa-gu-rŋi	+	+	+			m
1 incl plur	ŋV: - ru-	ŋa-gu-ru	+	+	+	m		

The gender features [PLUR], [MASC] and [FEM] are in complementary distribution and occur only in the [+aug] categories (for [+participant] categories). Later, I will present in section 3.2.1 an explicit account of the neutralizations and distribution of categories. For now, observe that the [-aug] categories have special forms, the 1sg ŋa-, pronoun ŋaya 'I,' and 1incl dual na-, pronoun nagawa: 'me and you (sg.).' The [+aug] forms have a common stem form to which are suffixed the three gender affixes -n̠i/-ni 'masculine,' (-r)ŋi 'feminine,' or -ru 'plural.'

Concentrating on the intransitive affixes only for now, I set up the following affix-feature correspondences (/V/ represents a harmonizing vowel):

(79) a. +I +you +aug ŋV:-
b. +I +you na-
c. +I +aug nV:-
d. +I ŋa-

(80) a. MASC -ni
b. FEM -ŋi
c. PLUR -ru

What is now of interest is that the augmented stem without the PLUR gender marker is interpreted as trial or dual, depending on the inherent number of the stem. That is to say, ŋV:-ni and ŋV:-ŋi are 'we excl.dual' masculine and feminine respectively and nV:-ni and nV:-ŋi are 'we incl. trial' masculine and feminine. The dual or trial number of these forms is *negatively* designated by the absence of the PLUR gender marker. (Because the PLUR gender marker is not present, these forms are able to show gender agreement, but this is a tangential fact.)

The interpretation of these forms as dual and trial must arise through functional inference. Because they are not plural gender, but at the same time not [-aug], they are interpreted conversationally as being augmented by one from the inherent number of the stem.

According to Heath, the trial number is optional. I take this to mean that the plural category may be used where only three referents are intended. This accords with the idea that the plural gender is vague with respect to number. It may be interpreted as one more than the highest more explicit number category, or as merely one more than the inherent number of the stem. Thus ŋV:-ru is understood as either at least 4 (one more than the trial form ŋV: -ni /-ŋi) or merely at least 3 (one more than the [-aug] form na-).

Because the trial and dual are the augmented categories without the explicit plural gender marker, they are less marked (for number) than the plural. There is no explicit mark for the dual or the trial (vis-à-vis the plural), nor are the dual and trial formed as composite pronouns as in Rembarrnga of Bamileke. As far as I am aware, Nunggubuyu is unique in designating duals and trials in this negative way. Consider Greenberg's Universal 35 (1963:112):

(81) There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and trial are almost never expressed only by zero.

The Nunggubuyu case is extremely unusual insofar as functional forces conspire to ensure that the morphologically unmarked forms are interpreted as dual or trial. These considerations are fundamental to an accurate analysis of multiple-argument agreement in the language, which I will discuss at length in chapter 3.

The same morphological pattern shows up in the 2nd person system as well:[12]

(82) Nunggubuyu 2nd person categories

category	prefix	pronoun (nominative)	[± you]	[± aug]	MASC	FEM	PLUR
2 sg	nun-/ba-	nagaŋ	+	-			
2 m dual	na-wV-ni-	nu-gu-ṉi	+	+	m		
2 f dual	na-wV-ŋi-	nu-gu-rŋi	+	+		m	
2 plur	na-wV-ru-	nu-gu-ru	+	+			m

Here again the [-aug] category has a special form, nun-/ba- prefix, varying according to tense, and pronoun nagaŋ 'you (sg).' The dual forms have the augmented stem followed by gender suffixes, but are not explicit marked as duals. The plural has the augmented stem followed by the plural gender marker.

To summarize, because plural is marked by a gender marker and because at the same time [±aug] is active in the system, the augmented forms not marked for plural are interpreted as inherently dual, or trial in the case of the 1st inclusive. Nunggubuyu presents the unusual case in which duals and trials are morphologically unmarked with respect to plurals, and the dual/trial interpretation is arrived at by functional inference alone. It is not particularly surprising then that dual and trial, being in a sense derived categories, should be particularly recessive when neutralizations level out the available combinations when the verb agrees with more than one argument. This more complex set of facts will be addressed in chapter 3.

---

[12]The segment /w/ here corresponds to Heath's /w_2/, which hardens to /b/ when following a nasal. There is another /w/ which hardens to /g/ following a nasal. See section 3.2 for further discussion.

2.2.7 The Person-Number Affixes of Sierra Popoluca

In this section I entertain and reject the claim that the Sierra Popoluca person-number system requires a special feature [±3]. I proposed in section 2.1.2 two Generalizations regarding the inclusion and exclusion of the role {3} in an argument's role set. These are repeated below:

(83) Generalization I: No person category specifically excludes {3}. (= 7A)
Generalization II: Unless {3} is the only permitted role, {3} is optional (=7B)

The effects of these generalizations were derived in part by the absence of any feature [±3] in Universal Grammar. The value [-3] would make {3} impossible (flouting Generalization I), and the value [+3] would make {3} obligatory (flouting Generalization II). The Sierra Popoluca person-number system presents a challenge to this analysis, since, *prima facie*, it seems to require [±3]. In fact, however, other evidence to be presented shows that this is (almost certainly) the result of misanalysis.

Sierra Popoluca is spoken in the state of Oaxaca, Mexico. As described by Foster & Foster (1948), this language possesses three 1st person nonsingular categories:

> The exclusive plural excludes the person or person addressed. The limited inclusive plural includes the speaker and the person or persons addressed, and excludes any others who may be present or referred to. The generalized inclusive plural includes the speaker, person or persons addressed, and any other person or persons present, or absent and referred to.
> (Foster & Foster 1948:19)

Sierra Popoluca, then, would appear to display one of Greenberg's "rare and doubtful" cases in which [-3] would need to be invoked. Specifically, the "limited inclusive plural" would have the matrix [+I +you -3]. Zwicky (1977) in fact takes Foster & Foster's account as the basis for postulating a feature [± III], and thereby predicting that some language (as yet undiscovered) should show an inclusive vs. exclusive 2nd person as well.

In this section, I show that [±3 / ±III] can be dispensed with entirely. The limited inclusive is in fact merely a misanalyzed 1st person dual.

2.2.7.1 The 1st Person Categories

The morphological reflexes of the 1st person categories are summarized below:

(84)	a. ta(n)-		1 limited inclusive plural
	b. a(n)-		1 singular
	c. a-	-taʔm	1 exclusive plural
	d. ta(n)-	-taʔm	1 general inclusive plural

Now, an examination of the other uses of the prefix ta(n)- shows that it functions as the mark of [+I +you], whereas the prefix a- is the mark of [+I (-you)]. What is of interest here, then, is the precise role of the suffix -taʔm.

I want to suggest that -taʔm serves as the mark of the feature [+augmented]. This gives the following pairings of features, affixes, and categories:[13]

(85)	a. ta(n)-		[+I +you -aug]	1 limited inclusive plural
	b. a-		[+I -you -aug]	1 singular
	c. a-	-taʔm	[+I -you +aug]	1 exclusive plural
	d. ta(n)-	-taʔm	[+I +you +aug]	1 general inclusive plural

Given the interpretive mechanisms I have proposed, ta(n)- without the suffix -taʔm, the 1st person limited inclusive (84a), should have a strictly dual interpretation, i.e. only the speaker and hearer are involved, since this is the minimal set fulfilling the conditions [+I +you]. The question of the correctness of this analysis then rests on Foster and Foster's assertion that (84a) refers to the speaker and person *or persons* addressed.

2.2.7.2 The Limited Inclusive as Dual

Subsequent research on Sierra Popoluca has supported my hypothesis. In particular, Ben Elson (p.c.) has informed me that his understanding of the category represented by (84a) is 1st person dual: speaker and hearer. The following examples illustrate.

[13] The t- prefix in (85a,d) may be segmentable as the realization of [+you] in the environment of [+I], but this is not entirely straightforward, so I retain the analysis as given.

(86) a. ta-nɨk-pa 'you and I go'
b. a-nɨk-pa 'I go'
c. a-nɨk-taʔm-pa 'I and those with me go'
d. ta-nɨk-taʔm-pa 'we all go'

The glosses in (86) support the interpretations in (85). Further illustrations of Elson's observations may be found in his work on Sierra Popoluca (1960a, 1960b, 1967).

Let us now consider the alternative hypothesis, namely that (94a/95a) is not in fact a 1st person dual, but rather a "limited" inclusive in Foster & Foster's sense, such that 3rd persons are excluded and more than one addressee may be involved. A formal problem vexes this analysis as well as the observational one just mentioned.

If (84a) is a "limited inclusive" then the first person categories must have the features given below:

(87)	a. ta(n)-		[+I +you -3]	1 limited inclusive plural
	b. a-		[+I -you -3 -aug]	1 singular
	c. a-	-taʔm	[+I -you +3 +aug]	1 exclusive plural
	d. ta(n)-	-taʔm	[+I +you +3 +aug]	1 general inclusive plural

In order to include the possibility of either one or more addressees being involved in (87a), this category can be neither [+aug], which would require at least two addressees, nor [-aug], which would require only one addressee. We are forced to conclude that if (87a) is a limited inclusive in Foster & Foster's sense, it has no value for [±aug]. Continuing with this line of reasoning, we next try to identify the conditions for the appearance of -taʔm From (87), it is appropriate to conclude that -taʔm is the marker of [+3 +aug].

To see this, suppose instead that -taʔm marks only [+aug] (= Elson's "plural"). Then we are forced to conclude that ta(n)- represents [+I +you -3] when -taʔm does not appear, but [+I +you +3] when -taʔm does appear. Yet this is really tantamount to saying that -taʔm represents [+3 +aug]. Yet this conclusion cannot be correct either when the 2nd person categories are examined:

(88)	a. mi-		2nd singular
	b. mi-	-taʔm	2nd plural

(89) a. mi-ñɨk-pa 'you go'
b. mi-ñɨk-taʔm-pa 'you all go'

If -ta?m really represents [+3 +aug], as the analysis in (87) would suggest, then the 2nd person plural (88b) and (89b) would have to be [+you +3 +aug]. This would mean that any use of the 2nd person plural in Sierra Popoluca would be "inclusive" in the sense that a 3rd party aside from addressees would be necessarily implicated. In other words, 2nd person plural would always have to mean at least 'you and him/her,' never just 'you and you.' Yet neither Foster & Foster nor Elson observe such a restriction.

An analysis of (84a) as a limited inclusive thus forces the peculiar result that the 2nd person plural is *always* "generalized," as long as we admit that the morphology of these forms is correlated with their interpretation through the intermediary of abstract features. All of these complications can be avoided by treating -ta?m as the marker of [+augmented] alone, as I suggested, and abandoning the limited inclusive interpretation.

2.2.7.3 Predictions of Usage

Certain subtleties of usage are now predicted. If a speaker is addressing several others, and wishes to refer to all present as a group, then Foster & Foster's description suggests that the limited inclusive (84a) would be used; my analysis on the contrary suggests that the so-called generalized inclusive, i.e. the augmented inclusive (84d) would be used. However, as Ben Elson has informed me (p.c.), "the plural does not seem to me to be a very important category in Sierra Popoluca. The plural morphemes are often omitted, once a plural participant has been established." In light of this last consideration, it is quite possible that where I predict that (84d) would be used, (84a) could be used in virtue of the omission of -ta?m for independent reasons. This may have led Foster & Foster to their conclusion. The decisive question would be: can (84d) be used if a speaker wishes to include more than one addressee but not include others? Such a usage would be impossible on Foster & Foster's account, but expected on my analysis.

Although this question cannot be answered definitively here, if I am correct in following Elson in labelling (84a) a 1st person dual, then the feature [±3/±III] can be dispensed with. This is a quite satisfying result, insofar as we can maintain the strong and apparently correct prediction that the possible non-composite person categories in human language can be limited to the four categories 1 incl, 1 excl, 2, and 3 (for languages which distinguish only 3 discourse roles).

2.3 Inherent Number and Agreement Class in Kiowa-Tanoan

In the following sections (2.3 to 2.3.3) I illustrate the intimate connection between inherent number and agreement class in the Kiowa-Tanoan languages. Of these, Jemez and Kiowa and Kiowa show the most complicated number system and will be discussed first in sections 2.3.1- 2.3.2. I then compare the Jemez-Kiowa system with that of Tewa and Tiwa, showing how these languages diverge from the proto-system, by relatively simple loss of rules, and, in the case of Tiwa, of the dual number for nouns.

All the Kiowa-Tanoan languages possess a rich system of inherent number for all nouns. Inherent number is not unknown in English: *scissors*, for example, although semantically singular, behaves morphosyntactically as a nonsingular. The Kiowa-Tanoan languages show a most expansive use of such inherent numbers. In Jemez, for example, animate nouns are inherently [+sg], but inanimates vary between being inherently [+augmented], inherently nondual, and inherently numberless (mass nouns). In addition to their inherent number, nouns are further specified with a syntactic number: basic or inverse. In the inverse form, the noun is interpreted as having the opposite number as it has inherently. For example, inherently singular nouns remain singular in the basic number, but become nonsingular in the inverse number. Both syntactic number, which I will call *number-class*, and inherent number interact in the verbal agreement system. This interaction will be the topic of the next several sections.

2.3.1 Jemez

The following account of Jemez is based on the description of Sprott (1991) and on the field work of Ken Hale, recorded in the late 1950s. Wherever significant differences occur in these two forms of Jemez speech, they will be duly noted.[14] Jemez is spoken today at Jemez Pueblo in northern New Mexico.

Jemez nouns divide into four classes based in their inherent number. On my analysis, Class I is inherently singular, class II inherently augmented, class III inherently nondual, and class IV are mass nouns with no number. As will be shown, the Jemez number system uses both [±sg] and [±aug] as in Hopi (section 2.2.2).

[14]In particular, I will generally be suppressing tone and stress. The tonal system of Jemez remains poorly understood and it seems best for now to leave off tone marks that are at best guesses to the phonologically significant contrasts of the language. I will also be suppressing the surface echo vowel which occurs finally after glottal stop.

There is but one form of number marking on the noun, a suffix -š. The distribution of this suffix in the four classes is shown below (Sprott 1991:53):

(90)

	I	II	III	IV
singular	-Ø	-š	-Ø	-Ø
dual	-š	-š	-š	-Ø
plural	-š	-Ø	-Ø	-Ø
inherent number	[+sg]	[+aug]	[+F]	---

We will consider the first three classes I, II and III, which have inherent number, and then turn to the variety of class IV nouns.

Consider first the class I nouns. When having no affix, these assume their inherent number, [+sg]: mɨ·sa 'cat,' ve·la 'man.' When suffixed by -š, they become *inverse* and assume the opposite value from their inherent number (and also show fronting of -a before -š).Thus mɨ·sæ-š means 'cats,' and ve·læ-š means 'men.' Such words are [-sg], and are therefore either dual or plural. The effect of the inverse number-class is to switch the inherent number specification to its opposite value.

Class II nouns work in exactly the same way, except that they are inherently [+aug]. Since only [-sg] arguments may be [+aug], they are redundantly [-sg] as well, but this value is not specified. The inverse number class turns class II nouns into [-aug] arguments, so that they become either singular [+sg -aug] or dual [-sg -aug]. Thus ta·sæ means 'dishes,' (more than two), whereas ta·sæ-š means 'one dish / two dishes.'

On the basis of classes I and II we have enough evidence to suppose that the Jemez number system, like the Hopi one, operates with the features [±sg] and [±aug]. Each of the four values that these feature provide defines a natural class in Jemez: [+sg] and [+aug] are the basic, inherent number of classes I and II, while [-sg] and [-aug] are the inverse numbers for these classes.

Class III is a bit more complex. In the basic form, class III nouns may be interpreted as [+sg], or [+aug] (more than 2). For example, to·pæ means 'one pencil, or more than two pencils.' When inverse, class III nouns are specifically dual: to·pæ-š 'two pencils.' Because the inherent number of class III is either [+sg] or [+aug], these nouns are lexically listed merely as [+F], where F is some number feature. If [+sg] is chosen as F, then [-aug] will be redundantly filled in; if [+aug] is chosen, then [-sg] will be redundantly filled in:

(91)	Class III	[±sg]	[±aug]
		+	(-)
		(-)	+

Class III nouns can also be analyzed as [αsg −αaug], if the lexical specification of number is left fully specified. However, the least redundant specification for this class is merely [+F], where F is either [±sg] or [±aug].

The inverse rule works straightforwardly for class III. When inverse, class III nouns are specifically dual. Just as in classes I and II, the inverse class switches the lexically specified value to its opposite. If F= [+sg] is lexically specified, then inverse is [-sg]. Along with the redundant value [-aug], this gives [-sg -aug] or dual. Similarly, if F = [+aug], then inverse gives [-aug]. The result is again [-sg -aug] or dual.

We can now formulate the number-switch inverse rule with maximal generality:

(92) Kiowa-Tanoan Number-Switch

N + INV	-->	N + INV	
[αF]		[−αF]	where F ranges over [±sg] and [±aug] and is lexically specified on N

The origin of the INV marker, which in Jemez is -š, but is −gɔ̀ in Kiowa, is unclear. One likely possibility is that INV is an independent syntactic head. Carstens (1991) has argued that the class prefixes of Bantu are independent number heads in syntax. The Kiowa-Tanoan inverse suffix is like a Bantu prefix in registering number; unlike the Bantu prefixes, however, the Kiowa-Tanoan inverse suffix interacts with inherent lexical number through the rule (92). Additionally, both semantic number and classificatory number (basic or inverse) play a role in determining agreement morphology, as is now detailed.

For simplicity, we confine our attention here to the intransitive prefix system. The transitive prefix system in some cases marks for more than one argument and presents a variety of neutralizations, some of which will be discussed in chapter 3.

There are four intransitive prefixes for 3rd person arguments in Jemez, as shown below:

(93)	A	Bi	Bii	C
	Ø-	į-	e-	ɨ-

Sprott labels the four classes A, Bi, Bii, and C. The two B classes neutralize in certain parts of the transitive agreement system. The chart below shows the distribution of agreement class by number and noun class (The behavior of class IV is complex and will be discussed later):

(94)

	I	II	III	
singular	A	Bii	A	[+sg -aug]
dual	Bi	Bi	Bi	[-sg -aug]
plural	Bii	C	C	[-sg +aug]

Both inherent number and number-class information determines which agreement class in (93) a given argument will require. For example, while class Bi is correlated simply with dual, class Bii is used in the singular for nouns of class I and in the plural for nouns of class II. These are exactly the nondual inverse categories, i.e. the inverse categories left over after the dual inverses have been assigned Bi. Moreover, C goes with the [+aug] (plural) class, and A with the [+sg] class, but only where these are not inverse and therefore take Bii. An ordered rule system will capture this descending hierarchy automatically:

(95)

a. INV, -sg -aug	Bi	ɨ̨l-
b. INV	Bii	e-
c. +aug	C	ul-
d. Elsewhere	A	Ø-

Several observations arise from this array of rules. The first is that both semantic number and classificatory number-class may condition the rules in (95). The first rule applies only to dual arguments. Since no noun is inherently dual, all dual arguments are also INV. Rule (94b) applies to the remaining INV categories, which happen to be class I plurals and class II singulars. The third and fourth rules distribute C and A in the elsewhere (basic) categories according to semantic number, C for [+aug] and A elsewhere.

It should be apparent that speakers of Jemez know which agreement affix to use not in virtue of knowing the simple table of correspondences in (94), but rather in virtue of knowing the rules in (95), which are ordered according to their relative priority of application. Furthermore, the fact that Sprott's categories Bi and Bii neutralize in the transitive agreement system follows not from the fact that they share some abstract agreement class, but rather that they share the classificatory feature INV, which is independently needed to derive Number-Switch as well as define the more specific classes picked out by rules (95a) and (95b). The letters that Sprott assigns to the agreement classes have no real grammatical value: they are only convenient labels specifying the classes defined by the rules in (95), which are defined in terms of the independently motivated number and number-class features.

Under certain circumstances, class II and class III nouns may take agreement A even when [+aug]. This seems to happen when the argument is understood as a collective plural, whereas the C agreement is a distributed reading. Sprott (1991:90) provides several examples, of which one is given here:

(96) a. ta·sæ Ø-šæ· [ta·sæ Ø-šæ]
dishes A- lie
'The dishes are here'

b. ta·sæ ɨ-šæ· [ta·sæ ɨǰæ]
dishes C-lie
'Dishes are here.'

The first example would be used if the dishes were taken as a unit; the second if the multiplicity of the dishes were being emphasized. Sprott remarks that the second example is "a bit 'more plural' for being less concentrated, more diffuse, more spread out, less identifiable as a collectivity or set. (p.90)"

Rule (95c), providing agreement category C, must therefore apply only if the argument is [+aug] and also not collective. If the noun is in fact collective, then rule (95c) will not apply and the final rule will provide A-agreement. This actually provides good evidence that the last rule is the Elsewhere rule. Under normal circumstances, A-agreement is limited to basic [+sg] categories (review (94)). If the rule supplying A-agreement positively identified the class as, say [+sg], then collective nouns of classes II and III, such as ta·sæ 'dishes' in (96b) would not take A-agreement, since they are still [+aug]. On the other hand, if the A-agreement rule is the Elsewhere rule, then it applies to all arguments remaining after the other rules, regardless of their feature-content. On the other hand, evidence to be reviewed shortly suggests that instead C is the default agreement class.

The class IV nouns present interesting properties. Heuristically, nouns designated class IV are those nouns which never take the inverse marker -š. Beyond this unifying property, however, class IV nouns are variegated. Sprott concludes that there are two sorts of class IV nouns: those which invariably have a single agreement class, and those which behave as class III for verbal agreement, but, unlike class III nouns do not take -š when semantically dual. I will now show how these facts can be explained naturally and without the puzzlement surrounding such questions as 'Is there really a class IV? Are some nouns of one class for nominal morphology and another for verbal morphology?'

The analysis to be presented here differs crucially from Sprott's in that I do not take class membership to be a primitive notion of Jemez grammar. We have already seen that classes I, II, and III are the labels which are attached to nouns which have various types of lexically specified inherent number. Consider the usage of the word *enemy* in British English. Both the phrases *the enemy is approaching* and *the enemy are approaching* are perfectly grammatical, although the latter is used only when the enemy consists of more than one individual. A similar circumstance holds in Jemez for those class IV nouns which are class III in the verbal morphology. Consider this pair of sentences (Sprott, p. 93; a colon [:] separates agent from patient in the gloss):

(97) a. wa·sa ta-mį̨
ice 1:A-see
'I saw ice'

b. wa·sa tɨ-mį̨
ice 1:C-see
'I saw ice(s)'

The second sentence would be used to refer to ice cubes. One could gloss it as 'I saw ices,' were it not for the English conditions against pluralizing mass nouns. What is of interest is that such nouns show distinction in the verbal agreement system while never showing inverse -š for number. Sprott identifies these as class III with respect to the nominal morphology. Yet this is really a conceptual error, as I will now show.

Returning to the rules in (95), observe that the first rule refers to the class of duals which happen also to be INV. Because all duals are INV, then rule (95) need not refer to INV class specifically. It will apply before (95b) in virtue of being more specific. Now consider which rules in (95) apply to class III nouns. When dual, these nouns are subject to (95a); when [+aug] (plural) they are subject to (95c), and otherwise, they are subject to (95d). What is characteristic of class III is that no rule of the verbal morphology must refer to INV. Now supposing that there is a class of nouns which are never INV, it follows automatically that those same rules will apply to this class as apply to class III. And that is exactly what happens. Nouns such as wa:sa 'ice,' never have an INV number-class, and therefore they are subject only to the verbal agreement rules which refer only to semantic number, and not to INV. These happen to be the same rules as apply to class III. This, however, is an accident arising from the fact that the INV of class III has its own special rule (95a) which refers specifically to a number (dual: [-sg -aug]).

One obvious result is that we should not expect to find nouns which have no inverse but which behave as class I or class II nouns. This is consistent with Sprott's findings: "I have not yet encountered any that pattern like Class I or Class II (p. 96)" On an analysis that takes the classes I-IV and A-D as primitives, this is a peculiar accidental gap. We should, other things being equal, expect to find class IV nouns patterning as I or II in the verbal morphology, if some behave as class III. On my analysis, where noun classes and agreement classes are not primitives, such a gap is not expected. Instead, the facts follow automatically and without any further stipulation.

Nouns such as wa·sa 'ice,' those that never take -š but behave as class III for the verbal morphology, have the following analysis. These nouns can never be inverse, and, furthermore, they have no inherent number. Recall that class I is inherently [+sg], class II inherently [+aug], and class III inherently [+F], where F is some number feature. There is a gap in this list, namely those nouns which have no inherent number specification: these are the class IV nouns. Although lexically numberless, they can have any cardinality. Their number features are freely supplied, and these then condition agreement. For example, wa·sa 'ice' is inherently numberless, but in (97b), the value [+aug] has been freely supplied to the noun and this conditions C-agreement on the verb. In (97a), the "mass" reading ('I saw ice, i.e. frozen on the ground'), the noun remains numberless. It therefore is interpreted as mass and is subject to the elsewhere agreement class, class A.

Sentences of type (97b) have another peculiarity. C-agreement class may refer to specifically noncollective plurals [+aug], but C-agreement has a number of other uses. First, it may be used with indefinite objects, such that we could also gloss (97b) as 'I ice-saw,' as well as 'I saw ices.' Nouns which never have distributive plural interpretations are thus interpreted in this fashion when used with C-agreement:

(98) hį? tɨ-mį
sand 1:C-see

'I sand-saw'

These distributed plural and generic readings are apparently two different semantic functions of C-agreement which have been morphologically neutralized in Jemez.

Second, C-agreement is used of arguments which are not overtly represented in the clause. Sprott (1992:115) gives the following pair:

(99) a. nɨ̨́ *July* tɨl-šæ·yo-hɨ̨ʔ (--> tɨjæ·yohɨ̨ʔɨ̨)
I July 1:C-give.birth.to-FUT

'I will give birth in July'

b. nɨ̨́ *July* šɨ tɑ-šæ·yo-hɨ̨ʔ (--> tɑšæ·yohɨ̨ʔɨ̨)
I July male 1:A-give.birth.to-FUT

'I will give birth to a boy in July.'

In the first sentence (99a), the transitive predicate 'give-birth-to' must have an empty object, which triggers the C-agreement. Where the object is overt and is [+sg], as in (99b), A-agreement appears. The appearance of C-agreement for empty NPs as in (99a) allows us to hypothesize that the 'I sand-saw' interpretation of (98) arises from a (string-vacuous) demotion of 'sand' from its object position, such that what is left as the object is an empty NP:

(100) a. nɨ̨́ *July* [$_{\mathrm{NP}}$e] tɨl-šæ·yo-hɨ̨ʔ (--> tɨjæ·yohɨ̨ʔɨ̨)
I July 1:C-give.birth.to-FUT

'I will give birth (to someone) in July'

b. hɨ̨ʔ [$_{\mathrm{NP}}$e] tɨ-mɨ̨
sand 1:C-see

'I ice-saw'

Similarly, C-agreement for the object is the regular means of expressing the reflexive (K. Hale, field notes):

(101) a. nɨ̨́ ʔɨ̨́ tɨ-tèle
I self 1:C-shot

b. ʔɨ̨́wá ʔɨ̨́ kɨ-tèle
you self 2:C-shot

On the basis of such facts, C-agreement might be analyzed as the default agreement class; Sprott makes this proposal explicit (1992:114).[15]

[15]Sprott asserts (1992:113): "Using the C-agreement prefixes for reflexives and reciprocals is not an unusual thing within the language family -- Kiowa does exactly the same thing-- ...' I would like to point out, however, that there is a systematic difference between Kiowa and Jemez in this regard. The agreement class which Jemez uses for reflexives is different from that used in Kiowa. As we have seen, Jemez uses reflexive objects with the C class, which also expresses *inanimate* distributive plurals. Kiowa, on the other hand, expresses reflexives with the same prefixes which express *human* plural objects. Inanimate

The problem with this proposal is that it requires that A-agreement be defined positively, since A cannot be the elsewhere class any longer. Sprott suggests that the A-class be defined as 'basic': in the terms presented here, this means that they are provided with the syntactic number-class BASIC. However, distributive plurals (i.e. the 'I saw ices' reading (97b)) are also BASIC inasmuch as they are not INV -- since they take neither INV agreement (class B) nor have the inverse marker -š. The A-agreement class must then be defined as BASIC *and* [+sg] to exclude the distributive plurals. This has two consequences. The first is that [+aug] nouns such as 'dishes' in (96a) must, when interpreted as a collective set, somehow be assigned the value [+sg] so as to permit A-agreement. Second, nouns interpreted as mass, i.e. (97a) ('I saw ice'), which take A-agreement, must also be [+sg], even though mass nouns are semantically numberless. On the other hand, if C is not the elsewhere class, then special provisions must assign [+aug] to reflexives and empty NPs to account for the phenomena in (101) and (100).

The two hypotheses can be exhibited below with the rules required:

(102)	C = Elsewhere		A =Elsewhere	
	BASIC [+sg] mass [+aug] collective	A	[+aug] distributed reflexive empty NP	C
	Elsewhere	C	Elsewhere	A

As can be seen above, neither the A nor C classes when positively identified is a simple natural class. Therefore, either analysis will have to include rules which assign mass/collectives to [+sg] or reflexives/empty NPs to [+aug]. I will henceforth adopt the

---

plural objects (the functional equivalent of the distributed plural C-class in Jemez) have a different series of prefixes (L. Watkins 1984:115):

Kiowa:

agent -->	1sg	2sg	2du	2pl	3sg	3du	3 human INV	3 INV
3 inanimate plur.object	gyàt	bàt	máǹ	bát̀	gyà	éǹ	gyá̀	ét
3 human plur./refl. obj.	dè	bè	mé	bé	èm	én	ém	ét

Jemez:

subject	1sg	2sg	2du	2pl	3sg	3du	---	3 INV
3 inan pl./refl. obj.	tɨl	kɨl	mɑ̨l	bɑl	ɨl	ɨ̨l		él
3 human pl./INV obj.	te	é	mɑ̨-pe	bɑ-pe	é	ɑ̨-pe		epe

In Jemez, human plurals are simply Inverse. If reflexive objects and human objects were neutralized as in Kiowa, then Jemez would treat reflexive objects as class B, i.e. INV objects. Thus, the two systems are not quite identical. Whether the specifically [+human] plural class is an innovation of Kiowa or an archaism is an open question.

assumption that A is the elsewhere class, although the other Sprott's hypothesis is also tenable.

I turn now to the second type of class IV noun identified by Sprott. These nouns always have the same agreement class in the verbal morphology, and, as class IV nouns, they never have INV -š. Examples include (class A): p'æ 'water,' ʔį 'blood,' gahwe 'coffee'; (class Bii) sį? 'rain,' cį·weš 'salt,' and (class C) cį 'grass.' Of these, I analyze the class A ones as inherently numberless like wa·sa 'ice,' but unlike 'ice,' these nouns can never have any number feature supplied to the them. In other words, they are strictly mass nouns and receive A-agreement by default. The feature [+mass] is incompatible with further semantic number, as is often the case cross-linguistically. The second type is also inherently and permanently numberless, but happens also to be inherently INV class. The last type is inherently [+aug] but cannot become INV.

What unites all the class IV nouns of Sprott's terminology is that none permit a syntactic INV marker, realized as -š, although some, e.g. sį? 'rain,' are inherently INV.[16] We can now tabulate the nouns according their lexical specification and two other criteria: whether they permit a syntactic inverse (i.e. not an inherent one, as in (103f)), and, if they are inherently numberless, whether they permit further specification of number:

(103) Jemez Noun Classes

	class	lexical value	permit INV?	further number?	example	
a.	I	[+sg]	yes		ve·la	'man'
b.	II	[+aug]	yes		ta·sæ	'dishes'
c.	III	[+F]	yes		to·pæ	'pencil'
d.	IV	Ø	no	yes	wa·sa	'ice'
e.	IV	Ø	no	no	p'æ	'water'
f.	IV	Ø, INV	no	no	sį?	'rain'
g.	IV	[+aug]	no		cį	'grass'

It is likely that nouns of type (103f) and (103g) are somewhat uncommon. It seems generally to be the case that nouns which are lexically numberless (103e,f) correspond to

[16]Alternatively, these nouns may require a syntactic inverse, but be lexically numberless. Sprott provides two examples of such nouns: sį? 'rain' and cį·we-š 'salt.' The latter always has the inverse marker, but is a mass noun. If these nouns are inherently INV and permit no syntactic inverse, then one has to say that 'salt' has the inverse suffix as part of its lexical representation. If these nouns require a syntactic inverse, then one has to say that 'rain' takes an inverse suffix which happens to be -Ø. I have no evidence on which to decide this matter now, although it may be significant that no such inherently INV nouns exist in Kiowa as far as I am aware.

mass nouns, although these appear to be subdivided between those which are strictly mass (103e), and those which permit a number distinction to appear on the verb. On my analysis, these nouns have number freely applied to them, but because they permit no syntactic INV marker, number is never reflected in nominal morphology, but rather only in the verbal agreement rules such as (95).

There are certain gaps and language-particular idiosyncrasies of the Jemez system. For example, why should there be no nouns which are inherently dual, i.e. [-sg -aug]? In the next section, I show that in Kiowa, certain nouns are in fact inherently dual, so that this is really an accidental gap in Jemez. Another gap is that there are no inherently [+sg] nouns which cannot have a syntactic INV. These nouns would be always singular, since INV could never make them nonsingular. One might suppose that proper nouns would be of this sort. Referring to animates, they are inherently [+sg] but they are never INV, if proper names cannot show number distinctions.

To summarize, the Jemez agreement system employs both number features [±sg] and [±aug] as well as a number-class INV, which I suggested inheres in a syntactic head which incorporates into the noun, like the Bantu prefix system (Carstens 1991). In addition, Jemez nouns are lexically specified in various ways with inherent number or class. They also are lexically specified as to whether they may have number freely applied (if inherently numberless) and whether they permit a syntactic inverse. The INV marker effects a Number-Switch rule (96), which reverses the inherent lexical number of the noun. The verbal agreement rules refer to both number and to INV. More generally, I have argued that the various noun classes and agreement classes (I,II, A, B, etc.) defined by Sprott (1991) are not primitives of the system. Instead, independently necessary inherent number and INV number-class along with a hierarchy of agreement rules generate the multiplicity of agreement patterns quite directly without these primitives.

2.2.2 Kiowa

In this section I present an analysis of the Kiowa number system based on L. Watkins' (1984) study of the language. Kiowa is spoken today in southwestern Oklahoma mostly in Caddo, Kiowa and Comanche counties, although around 1700 the Kiowas lived east of the Black Hills of South Dakota (L. Watkins 1984:1), and migrated through the Plains region after the introduction of the horse to their culture.

It will be shown that Kiowa uses the same primitives and the same rule system as does Jemez but differs in what are possible lexically-specified numbers for nouns, depending crucially on the role of dual in the inherent number scheme. I have tried to formalize and expand the account provided by L. Watkins (pp. 78-86) and will also draw comparisons to Jemez as I understand them. I propose that Kiowa and Jemez split according the role accorded the dual. Later, in section 2.2.3, I show that Tewa agrees with Jemez, and therefore Kiowa may be the innovating member of the Kiowa-Jemez split.

Since Hale (1962) conclusively demonstrated the relatedness of Kiowa and Jemez on the basis of phonological correspondences, it has been presumed that both languages descend from a common source, as was first suspected by J.P. Harrington (1928). Even a cursory examination of the formal parallels in the two number systems, which can now be done thanks to recent work on both languages, substantiates this claim in the area of morphology.

Clearly, the correct analysis of the number systems in Kiowa and the Tanoan languages, including Jemez, is a first step towards reconstruction of the number system in this proto-language. On the surface, the two languages seem to have somewhat different variants of the same type of system: both have inherent number and inverses. But the agreement classes do not correspond exactly. As long as reconstructive efforts continue to operate with these classes as primitive grammatical notions, I believe they will be fruitless. However, once the actual lexical content of the various classes is identified, certain correspondences emerge.

As in Jemez, nouns in Kiowa have an inherent number, but the possible inherent numbers are different from those in Jemez. The following array is taken from L. Watkins (1984:78:ff), and duplicates the classification proposed by Wonderly, Gibson & Kirk (1954) and further discussed in Merrifield (1959):

(104)		I	II	III	IV
	singular	-Ø	-gɔ́	-gɔ́	-Ø
	dual	-Ø	-Ø	-Ø	-Ø
	plural	-gɔ́	-Ø	-gɔ́	-Ø
inherent number		[-aug]	[-sg]	[-sg -aug]	varies

The Kiowa inverse marker is -gɔ́ or any of its several allomorphs. As in Jemez, there appear to be four main classes of nouns based on their inherent number. The class I nouns are generally animates, although there are a number of zoologically inanimate nouns such as k'ɔ́dál 'wheel, wagon, vehicle' kʰɔ̃· 'car (<Eng.),' ɔ́lp'ɔ̨́ɔ̨· 'soap,' etc. (L. Watkins 1984:83). Classes II and III are inanimates, just as in Jemez.[17] However, the inherent number of class I in Kiowa is [-aug], that is, nonplural, whereas recall that in Jemez class I (animates) are [+sg]. The difference is that a Jemez animate noun will be strictly singular when having no inverse suffix; a Kiowa animate class I will be either singular or dual. Similarly, whereas class II in Jemez is [+aug], class II in Kiowa is [-sg]. Again, the crucial case is the dual: Jemez class II duals are inverse, whereas Kiowa class II duals are not inverse. Finally, class III in Kiowa is inherently dual, that is [-sg -aug], whereas class III in Jemez is exactly the opposite: nondual.

Examples of Kiowa nouns of classes I, II and III are provided below:

(105) a. cę̂· 'horse, 2 horses'
b. cę̂·-gɔ̀ 'horses'

(106) a. tʰǫ́·sè 'bones (2 or more)'
b. tʰǫ́·sè-gɔ̀ 'one bone'

(107) a. k'ɔ̃́n '2 tomatoes'
b. k'ɔ̃́n-dɔ̀ 'one tomato, more than two tomatoes'

In (105) and (106), the inverse suffix -gɔ́ undergoes regular tone lowering; in (107) it appears as allomorph -dɔ̀ after /n/.

The most salient difference between the two number systems is that in Jemez no noun is inherently dual, so that all duals must be inverses; in Kiowa, quite the contrary, no

[17]Actually, motility appears to be more relevant than animacy to this distinction. While there is a great deal of redundancy in the class system, that is to say that the semantic properties of a noun determine in large part its class membership, this is not always true or at least obvious by European conceptions of classification. I leave this complex topic open, since it is not directly relevant to the formal properties of the morphology.

noun is inherently nondual, so that all duals are basic. Aside from this difference, the two systems are otherwise remarkably similar.

On the basis of the inherent numbers of classes I and II, we can set up the following functional correspondences between Jemez and Kiowa:

(108) Jem. [+sg] ≈ K. [−aug]
Jem. [−sg] ≈ K. [+aug]

It is unclear which value in these functional correspondences represents the more archaic state, that is to say, whether the Kiowa system or the Jemez system is the proto-system, or perhaps a variant of both. What one can conclude, however, is that the dual number is the pivot in the change from one system to the other.

The syntactic INV -gɔ́ has the same number-switching effect in Kiowa as it does in Jemez, i.e., Kiowa is also subject to the Kiowa-Tanoan Number-Switch rule proposed in the last section . Inherently [-aug] nouns of class I become [+aug] when inverse (105); inherently [-sg] nouns, class II, become [+sg] (106); and inherently dual nouns [-sg -aug] become nondual, either [+sg -aug] or [-sg +aug], depending upon which inherent number feature undergoes the switch (107) (Both features cannot undergo the switch since the resulting matrix would be semantically incoherent: *[+sg +aug].)

Formally, then, the Kiowa system has the very same primitives and the very same Number-Switch Rule as does the Jemez system, but it differs with regard to what are the possible inherent numbers. Specifically, the Jemez inherent numbers are all defined in terms of the [+] values of the features: [+sg], [+aug], and [+F], where F is [sg] or [aug]. In Kiowa, quite the opposite, the inherent numbers are defined in terms of the negative values of the features: [-aug], [-sg], and [-sg -aug].

When we turn to consider the intransitive prefix agreement rules, there is a striking formal similarity with Jemez. As in Jemez, there are four intransitive agreement prefixes, distributed as below (I delay discussing class IV):

(109)

	I	II	III
singular	Ø-	è-	è-
dual	ȩ̀-	ȩ̀-	ȩ̀-
plural	è-	gyà-	è-
inherent number	[-aug]	[-sg]	[-sg -aug]

Just as in Jemez, the dual categories have a special prefix, ę̀-. INV classes have a special prefix è- which appears regardless of the semantic number of the argument. For example, class II inverses are singular and take è- whereas class I inverses are plural and so also take è-. This shows that agreement rules must refer in some cases to semantic number but in others to only classificatory number (INV). Finally, as in Jemez, the non-inverse (basic) plurals [+aug] have a special prefix gyà-, and Ø- is the elsewhere prefix, *de facto* appearing in the singular of class I:

(110)

	Kiowa	Jemez functional equivalent
a. -sg -aug	ę̀-	įl-
b. INV	è-	e-
c. +aug [–collective]	gyà-	ɨl-
d. Elsewhere	Ø-	Ø-

While it seems plausible to etymologically relate K. ę̀- with Jem. įl- and the inverse markers in both languages, the relation of K. gyà- and Jem. ɨl- is not obvious. L.Watkins derives gyà- from underlying *ìà-d, where the final -d marks a 'nonsingular object.' For reasons which are unclear, apparently final -d metathesizes with the vocalic syllable nucleus giving dìà-; by regular phonological rules d- becomes a velar g- before i. The underlying form *ìà-d is potentially cognate with Jem. ɨl-, since all syllable final consonants in Jem. are reduced to -l.

The Kiowa agreement system has a further complication: human plurals comprise a special class. Because human nouns belong in class I, the plural of a human noun will be syntactically INV; we expect then the intransitive prefix è- by rule (110b). However, human plurals have the special prefix á-.[18] The following rules/affixes incorporate this complication:

(111)

a. -sg -aug	ę̀-
b. INV +human	á-
c. INV	è-
d. +aug [–collective]	gyà-
e. Elsewhere	Ø-

---

[18]L. Watkins indicates that the use of this category is in actual practice limited to 'adult members of one's own tribe. (p. 113). Thus [+human] may be an insufficient determinant of the conditions of use for this category.

The formal equivalence of the agreement rules in Kiowa and Jemez (with the exception of the specific human INV of Kiowa) leads to the following diachronic picture of the Jemez-Kiowa split. While the inherent number system of the two-languages is different in virtue of the correspondences (108), observe that the distribution of agreement (at least in classes I and II in both languages is the same). This congruence is guaranteed because the dual has its own special agreement rule in both languages (93a) and (110a), and the dual category is the pivot in the Jemez-Kiowa split. Because the dual can be unambiguously identified through verbal agreement morphology in all cases in both languages, then in terms of nominal morphology, the dual, if originally always INV as in Jemez, was free to lose its INV marking with no threat of ambiguity. Or, if the change happened in the reverse direction, and duals were originally not INV as in Kiowa, then they could acquire INV marking in the nominal system without there having to be any changes in the verbal agreement rules.

Just as in Jemez, Kiowa nouns of class II when plural may be understood as collective, as was shown originally by Merrifield (1959). In such a case rule (110c) does not apply to them, and they therefore end up in elsewhere class (104d) when plural and when singular. However, it appears to be the case in Kiowa that for any given noun of class II, it is a lexical property whether it forms a collective plural or a distributed plural (L. Watkins, p. 86-87), and the semantic basis for this difference is apparently arbitrary. We may suppose, then, that those nouns which form only collective plurals are marked lexically as being incompatible with some feature [-collective] or [+distributed].

Class III nouns, as in Jemez, may also be collective. When referring to "three or more separate collections of a single type (L.Watkins, p. 88)," i.e. as the plural of the collective, they agree as if they were singulars of class I, that is neither inverse nor augmented. Since class III is inverse when plural, one would otherwise expect the plural of the collective to be inverse, but in fact it agrees as a singular:

(112) hɔ́ndé ɔ́l bɔ́-sɔ̨́·mí
what.INDEF hair 2plPAT.sgOBJ-interesting

'What interesting (kinds of) hair you-all have.' (L. Watkins ex. 26d, p. 89)

The verb shows the prefix bɔ́ indicating a 2nd person plural patient (the people who have the hair), and a singular object (the different kinds of hair). If the object were inverse, the expected prefix would be bɔ́t- (L.Watkins, p. 116). This suggests (to me) that INV and a collective reading are mutually exclusive, and that as a result, the elsewhere 2pl patient bɔ́ prefix must be used.

Turning to the class IV nouns, L. Watkins identifies three subclasses which find remarkable parallel in Jemez. These subclasses are shown below with their agreement prefixes:

(113)

	IVa	IVb	IVc
singular	Ø-	Ø-	gyà-
dual	ę̀-	ę̀-	gyà-
plural	gyà-	Ø-	gyà-
inherent number	Ø	Ø	[+aug]

As in Jemez, all nouns labelled class IV share the property that they cannot have a syntactic INV marker.[19] Nouns of subclass IVa have no inherent number, but semantic number may be supplied to them freely. Having no inverse, they are subject only to rules (110a,c,d) according to the semantic number which they receive. Examples include c'ó· 'rock, stone,' kḯ 'meat,' t'ɔ́á· 'earring.'

Such nouns are formally parallel to the Jemez inherently numberless class (96d), for example Jem. wa·sa 'ice.' However, recall that I argued in the last section that the fact that Jemez inherently numberless nouns behave as class III for verbal agreement was merely an accident of the agreement rules: it turned out that nouns having no INV and inherently nondual nouns in Jemez are subject to the same set of agreement rules. If there *were* an inherently nondual class in Kiowa, the same phenomenon would occur (because the agreement rules are formally parallel), but in fact there is not. The inherently numberless nouns in Kiowa do not behave like any other class, as is shown by the fact that no other class of nouns has the prefixes Ø-/ę̀-/gyà- in the singular, dual, and plural. This shows that L. Watkins is correct in saying that these nouns "are signaled in the prefixes according to actual number (p.89)." The very same rules apply in Jemez to its numberless nouns (96d).

Nouns of class IVb are identical to IVa-- they have no inherent number-- except they cannot form distributed plurals. As L. Watkins observes, these nouns are formally similar to those nouns of class II which cannot form distributed plurals. When they are plural [+aug] they nevertheless will not undergo rule (110c) and so surface with the elsewhere Ø- prefix. Examples include hɔ́ɔn'road' and sé·có· 'pond.'

[19] L. Watkins also cites (p.84) the noun t'áp 'deer,' which, like nouns of class IV, never has an overt inverse marker. However, when semantically plural, it agrees as an inverse. One possible analysis is that this noun has a special -Ø INV allomorph.

I suggest that nouns of class IVb correspond abstractly to the the strictly mass nouns of Jemez, e.g. p'æ 'water.' The Jemez strictly mass nouns *behave* as if [+sg] in the verbal agreement system, always showing the elsewhere Ø- prefix in intransitives, for example. The Kiowa IVb nouns, in virtue of forming only collective plurals, never show the [+aug] morphology gyà- and therefore *behave* morphologically as if they were only [-aug]. This finds a parallel in the inherent numbers of class I in the two languages. By the correspondence in (108), the value [+sg] in Jemez corresponds to [-aug] in Kiowa.

Finally, class IVc is inherently [+aug], like English *scissors*. Regardless of their actual number, they undergo rule (110c) and receive the [+aug] agreement prefix gyà-. Thus the word tó· when meaning 'teepee,' always takes a plural verb, regardless of whether one or more teepees are being referred to:

(114) tó· gyà-sɔ̀l
teepee [+aug] - be.set.[+aug]

'There is/are one/two/several teepee(s) standing there.' (L. Watkins' ex.29, p.90)

Nouns of this class correspond to the inherently [+aug] nouns of Jemez, e.g. cɨ 'grass,' (g).

We can now classify the Kiowa nouns according to their inherent number and two other criteria: whether they permit INV and whether their [+aug] forms are distributed or collective.

(115) Kiowa noun classes

	class	lexical value	permit INV?	[+aug] form?	Example	Jemez equivalent
a.	I	[-aug]	yes	distributed	cę̂· '(2) horse (s)'	(96a)
b.	IIa	[-sg]	yes	distributed	dɔ́ál 'buckets'	(96b)
c.	IIb	[-sg]	yes	collective	tól 'pegs, stakes'	id.
d.	III	[-sg -aug]	yes	either	k'ɔ̂n '2 tomatoes'	(96c)
e.	IVa	Ø	no	distributed	c'ó· 'rock, stone'	(96d)
f.	IVb	Ø	no	collective	hóɔ̀n 'road'	(96e)
g.	IVc	[+aug]	no	distributed	k'ólpʰą· 'necklace'	(96g)

I have not come across any citations of inherently INV nouns in Kiowa. These would have to show the intransitive agreement marker è- regardless of actual semantic number, and would correspond to Jemez sɨʔ 'rain.'

To summarize, the Kiowa and Jemez number systems are remarkably similar. They differ crucially in the treatment of the dual number in the inherent number system. In Jemez, duals are always INV, because no noun is inherently dual. In Kiowa, all duals are basic and no noun is inherently nondual. The agreement rules (at least for intransitive prefixes) have remained formally identical in both languages and are surely archaic. Because the dual number has its own special prefix rule, moving the dual from always basic to always inverse in the nominal morphology (or vice versa) was a small change, effected by means of the feature correspondences in (108). Taking into account these correspondences, almost all other details of the two systems of inherent number are identical, including the numberless nouns and the special uses of collectives for inanimate plurals. This surprising correspondence suggests that the Kiowa-Jemez system preserves the proto-Tanoan number system most fully. In the next section, I present analyses of Tiwa and Tewa, tracing their divergence from the archaic system.

2.2.4 Tiwa and Tewa

In this section I present analyses of the number systems in three other Tanoan languages, Southern Tiwa, Northern Tiwa (Taos), and Rio Grande Tewa, all spoken in New Mexico. Southern Tiwa is spoken at Isleta and Sandia pueblos; the data presented here were collected at Isleta. Taos is spoken at Taos pueblo, two miles north of the town of Taos (Trager 1946: 184). Rio Grande Tewa is spoken in five pueblos in New Mexico: San Juan, Santa Clara, San Idelfonso, Nambe and Tesuque (Speirs 1966:4).

The Tiwa systems present a considerable simplification from the Kiowa-Jemez systems presented in the last two sections. Southern Tiwa, in particular, is (apparently) the least complicated. According to Allen, Frantz, Gardiner & Perlmutter (1990; henceforth AFGP), nouns in Southern Tiwa fall into three classes. Below, the three classes are designated by Roman numerals I-III and the entries in the chart show the corresponding intransitive agreement affix. I have provided in the bottom line the inherent number of each class on my analysis:

(116)

	I	II	III
singular	Ø-	Ø-	i-
plural	i-	u-	u-
inherent number	[+sg]	Ø	[-sg]

The distribution of these three intransitive prefixes defines three agreement classes which are also relevant for multiple-argument agreement. Rosen (1990:672) provides examples of nouns in the three classes below. The agreement class letter used by AFGP is given in parentheses:

(117)

	I	II	III
singular	musa-de (A)	shut (A)	keuap (B)
plural	musa-n (B)	shut (C)	keuap (C)
	'cat'	'shirt'	'shoe'

Observe that in (117) the class II and III nouns are invariant, but nouns of class I show overt number affixes. There are allomorphs number affixes besides -de/-n for class I, but I will for simplicity confine discussion only to these. Thus, So. Tiwa presents a difference from the Kiowa-Jemez system: here not only may inverse nouns show an overt syntactic affix for number, but so also may basic nouns of class I. We return to this difference in the discussion of Northern Tiwa (Taos).[20]

The distribution of agreement affixes in (116) follows on the assumption that class I is inverse when plural and class III is inverse when singular. Class II is never inverse, but has agreement assigned according to (freely supplied) semantic number. The distribution of inverse and basic is given in (118) and the agreement rules in (118):

(118)

	I	II	III
singular	basic	--	INV
plural	INV	--	basic
Inherent number	[+sg]	Ø	[-sg]

[20]Whether Kiowa-Tanoan lost the overt suffixes for the Basic number or whether these were innovated in Tiwa-Tewa remains an open question. The traditional view (following Trager 1946) is that Tiwa is morphologically conservative (as it is phonologically); but this need not apply to all aspects of morphology. Archaism of basic *-nV would be strengthened if traces of this suffix could be found in Kiowa-Tanoan. I leave this question open.

(119)			
	a. INV	i-	(C)
	b. -sg	u-	(B)
	c. Elsewhere	Ø-	(A)

As can be seen in (118), the syntactic INV switches the number of the noun to the opposite value from that which it has inherently. The Tiwa system is considerably less complex than the Kiowa or Jemez systems, since it lacks a dual for third person arguments.

The system in Northern Tiwa (Taos) is similar to that of Southern Tiwa, except that all nouns show overt affixes both for basic and for inverse number (Trager 1946, 1954). Trager divided the nouns into three classes ("genders") each having two declensions, depending on the distribution of number markers. It can be seen that Taos I, II, III are equal to So. Tiwa I, III, II respectively.[21] The distribution of number affixes and agreement classes is shown below:[22]

(120) Northern Tiwa (Taos) nominal classes (Trager 1946)

	I		II		III	
	1	2	3	4	5	6
[+sg]	-na	-na	-ną	-nemą	-na	-ne
[-sg]	-ną	-nemą	-ne	-ne	-ne	-ne
[+sg]	A		B		A	C
[-sg]	B		C		C	C
Example	c̦ up'a- 'judge'	k'a- 'mother'	t''awa- 'wheel'	c'ine- 'eye'	kʷ'o- 'ax'	c''o- 'liver'

Comparing the Kiowa, Jemez, and Southern Tiwa intransitive prefixes, we find that Southern Tiwa has merely lost the special dual rule of Kiowa and Jemez, but preserved all the other rules intact, with the reanalysis of u- as [-sg] rather than [+aug]:

[21]I am retaining the classification of the original descriptions to avoid any further confusion. It is of course unfortunate that functionally and/or etymologically analogous classes should have different numbers. Compare the use of "circumflex" accent in Lithuanian and Greek. Tanoan studies are not immune to this problem of nomenclature.
[22]The marks /̦ / and /' / in Trager's transcription represent pitch-accents ("prosodic phonemes") with a complicated distribution (Trager 1946:189-90).

(121)		Kiowa	Jemez	So. Tiwa	
	a. -sg -aug	ę̀-	įl-	none	
	b. INV	è-	e-	i-	
	c. +aug	gyà-	ɨl-	u-	[-sg]
	d. Elsewhere	Ø-	Ø-	Ø-	

To summarize the discussion of Tiwa, the innovation of the Tiwa branch is the loss of dual number entirely from the system of inherent number. Duals always fall together with plurals in the number-affix system and with respect to verbal agreement. Additionally, Tiwa shows overt affixes both for basic and for inverse number. This phenomenon is limited to class I in Southern Tiwa, but has spread throughout the nominal system in Northern Tiwa (Taos), such that all nouns when standing absolutely have a number-class affix, as shown in (120). In this way, the Taos system resembles even more closely the Bantu prefixing system, since all number affixes are overt.

The number-agreement system of Rio Grande Tewa (Speirs 1966) looks on the surface rather unlike the Jemez-Kiowa-Tiwa system. The agreement prefixes (for third person arguments) are distributed according to the parameters of animacy and semantic number, without reference to the INV number-class, as follows:

(122)		
	a. -sg -aug	da-
	b. +aug +animate	di-
	c. Elsewhere	na-

Dividing nouns into animates and inanimates gives the following distribution of the affixes:

(123)		animate	inanimate
	singular	na-	na-
	dual	da-	da-
	plural	di-	na-

Etymologically, it seems that the di- of Rio Grande Tewa may be related to Kiowa gyà-, which derives from underlying d-ià, with d --> g by a rule switching dentals to velars before i (L. Watkins 1984:43).

Although the distribution of verbal agreement need not make reference to INV, evidence that INV continues to function as a syntactic class can be discerned in the nominal and demonstrative inflections.

Speirs (1966) divides the Tewa nouns into seven classes depending upon their inflection for number. His classification is given below; on the last line I indicate what I propose as the inherent number for each of Speirs' classes:[23]

(124)	A_1	A_2	A_2	I-dis_1	I-dis_2	I-ag_1	I-ag_2
singular	-Ø	-Ø	-Ø	-Ø	-Ø	-Ø	-nɨn
dual	-Ø	-ʔɨn	-n	-Ø	-n	-Ø	-nɨn
plural	-Ø	-ʔɨn	-n	-Ø	-Ø	-Ø	-di
example	cé	kóʔô·	xa'wi-	kʔu·	pu̥wa̜eʔi	te·	ta̜ʔ-
gloss	'dog'	'aunt'	'singer'	'rock'	'mirror'	'tree'	'book'
Inherent number	Ø	[+sg]	[+sg]	Ø	[+F]	Ø	[+aug]

A stands for animate and I for inanimate; 'dis' means the plural form has a distributed (noncollective) reading, and 'ag' means the noun has an aggregate (or collective) reading. The two A_2 classes are distinguished by the allomorph for [-sg] which they take.

Observe that for each class, A, I-dis, and I-ag, there is a subclass of nouns which take no number affixes. These are the analogues of the class IV invariant nouns in Kiowa and Jemez which permit no overt INV marker.

The overtly suffixed forms A_2, I-dis_2 and I-ag_2, on the other hand, have inherent number and correspond to Kiowa-Jemez classes I, III, and II respectively.

The distribution of the allomorphs of the demonstrative ʔoʔi 'this' is of interest in showing what the inherent numbers of the Tewa noun classes are. The distribution of the demonstrative by class is given below:

[23]Speirs' apostrophe /'/ indicates laryngealization of preceding vocoid.

(125) ʔoʔi 'this'

	A	I-dis	I-ag
singular	ʔoʔi	ʔoʔi	ʔoʔi-n
dual	ʔoʔi-n	ʔoʔi-n	ʔoʔi-n
plural	ʔoʔi-n [24]	ʔoʔi	ʔoʔi

I propose that the unaffixed demonstrative ʔoʔi goes with the basic forms. This means that the A class is inherently [+sg], the I-dis class inherently nondual ([+F], where F =[sg] or [aug]), and the I-ag class is inherently [+aug]. In this way, the three classes agree with Jemez classes I, III and II in their inherent numbers. All the inverse numbers have the allomorph ʔoʔi-n. The rules for the demonstrative can be given:

(126) a. INV -n
b. Elsewhere -Ø

One can also see evidence of a basic-inverse distinction in the nominal affixes in (124). For example, while xa'wi- 'singer' has -n in all [-sg] categories, inanimate noun pųwąeʔi 'mirror,' has -n only in the dual. The latter noun is inherently nondual, like Jemez class III, and therefore becomes only dual when inverse -n is affixed. But xa'wi- 'singer,' being inherently [+sg], becomes [-sg], i.e. either dual or plural, when affixed.

Similarly, in the I-ag$_2$ class, observe that the affixes -di [+aug] and -nin [-aug] are correlated with the affixes for class I in Southern Tiwa in (117), where -de is basic and -n is inverse. This suggests again that Tewa -di [+aug] is a basic number-class suffix and -nin is an inverse number-class suffix. This in turn allows us to hypothesize that the I-ag$_2$ class is inherently [+aug] in Tewa; and this hypothesis is supported by the behavior of the agreeing demonstrative as already discussed.

Tewa shows that same distribution of verbal suppletion as in Hopi (section 2.2.3). A subset of verbs and adjectives show suppletion, but only in the [+aug] (plural). The dual and singular share the other verb form and comprise the [-aug] categories (Speirs 1966:154):

[24]Final /-n/ of this form assimilates to a following /p/ by regular rules of sandhi (cf. Speirs, p. 49) In Speirs' examples, the demonstrative is followed by the number *po:ye* 'three,' so final /n/ becomes [m]. Thanks to Ken Hale for pointing this out

(127) a. ?o?i wî' cé he·?i na-ket?ą́
that one dog big 3-fell

b. ?o?i-n wíye cé ha?ennin da-ket?ą́
those-INV two dog big.nonsg dual-fell

c. ?o?i-m [25] po·ye cé ha?ennin di-yemu
those-INV three dog big.nonsg aug.anim-fell

The verb 'to fall' shows a suppletive allomorph -yemu for the [+aug] number in (127c). Its [-aug] allomorph -ket?ą́ appears in the dual and singular in (127a,b). However, the distribution of the adjective he·?i is conditioned by [±sg]: he·?i appears with the [+sg] noun in (127a) but ha?ennin with the [-sg] nouns in (127b,c). Again, the proper distribution of forms follows directly in a number system in which dual is negatively designated as [-sg -aug].

In sum, despite certain morphological complexities, such as the various allomorphs of INV, Tewa actually presents a number-system surprisingly similar to that of Jemez. In particular, nouns are either inherently numberless and show no INV (A_1, I-dis$_1$, I-ag$_1$) or have an inherent number [+sg], nondual, or [+aug] as in Jemez. The innovation of Tewa is that verbal agreement affixes (e.g. intransitive affixes in 122) are never conditioned by INV, but instead reflect only semantic number and animacy. The INV class continues to be evident in the distribution of nominal affixes and in the distribution of the demonstrative.

From an historical point of view, we may say that Tewa is moving away from an inverse number system to a system based purely on semantic number and animacy. The first step in this move is the loss of INV as a class relevant for verbal agreement. Tewa has no analogue to the INV prefix which is represented in Kiowa as è-, Jemez e- and Tiwa i-. However, the fact that the inherent numbers of Tewa agree with those of Jemez, and not those of Kiowa, suggests that Tewa and Jemez should be grouped together with respect to the Kiowa-Jemez split discussed in the last section:

(128) Dual always inverse: Jemez, Tewa (positive-valued inherent number)
Dual never inverse: Kiowa (negative-valued inherent number)
No dual in nouns: Southern Tiwa, Taos

[25] See fn. 24.

Finally, the number-class system of Tiwa is essentially that of Kiowa-Jemez except that Tiwa has lost the dual number for nouns. In so doing, it lost the special dual prefix which appears in Jem. as į̀l-, Kiowa ę̀- and Tewa da-.

In terms of archaism, Kiowa and Jemez preserve the oldest and most complex number system of Kiowa-Tanoan, although vestiges of the earlier stage are apparent in Tewa as well.

From a more general theoretical point of view, the Kiowa-Tanoan system shows a most expansive use of inherent number as is familiar in more restricted form from the *pluralia tantum* in European languages. I suggested that the INV suffix is really a number-class head in syntax like the prefix Number head in Bantu as analyzed by Carstens 1991. Taos presents the most overt evidence of this system of number-affixes, since all nouns when standing absolutely must have a number-class affix.

The Kiowa-Tanoan INV, in addition to expressing number-class, has the additional feature-switching effect as formulated in rule (92). This rule requires the α–notation. I provided examples of α–rules for person features in the discussion of Mam and Sursurunga in sections 2.1.6 and 2.1.9: the Kiowa-Tanoan Number-Switch provides evidence that number-feature α–rules are possible and necessary as well.

Finally, the functional correspondence between [±sg] and [±aug] allows a simple expression of the difference between the Kiowa and Jemez systems. This, and the groupings of basic vs. inverse numbers in the two languages, as well as the distribution of suppletive verbs and adjectives in Tewa, provides solid support for the proposed number feature system in which duals may be negatively designated [-sg -aug], as proposed for Hopi in section 2.2.2.

2.4 Summary

The number features [±sg] and [±aug] are the only bivalent number features. The former is defined simply as singleton vs. nonsingleton set of entities; the latter in terms of the proper subset relation. Even exceedingly complex pronominal systems, such as are found in the composite pronouns of Bamileke, are built up by the accumulation of the person-number categories definable by means of these features and the person features [±I] [±you] and [±prt].

Beyond these, languages may use explicit attributes of cardinality such as [dual] and [trial]. However, I showed that the *de facto* dual number arises in a number of ways, either by explicit mark [dual], by negative designation [-sg -aug], by inherent number, or by functional inference. Most typically, 1st person inclusives are inherently dual and therefore need not be explicitly dual. Nouns may be inherently dual, as in the class III nouns of Kiowa. I also showed that functional inference supplies number in complex pronominal systems, in particular in Nunggubuyu, a [PLURAL] gender feature may interact with inherent number to give morphologically unmarked duals and even trials in the 1st inclusive.

Finally, the Kiowa-Tanoan languages show a syntactic number-class affix, INV, which switches the inherent number of an argument to its opposite value. This affix is formally parallel to the Bantu class prefixes, except that it performs the additional switching effect. Such a rule provides evidence that α-notation is necessary for number features, as well as to register agreeing or disagreeing person features as in Mam.

The analyses of Nunggubuyu and Kiowa-Tanoan will now form the bases for a more detailed consideration of the multiple-argument agreement of these languages in the next chapter.

CHAPTER 3: Problems in Multiple Argument Agreement

3.0 Introduction

Languages whose verbs show agreement with more than one argument present special complications for any theory of inflection. The complications fall under three broad and interactive headings, which are not specific to multiple-argument agreement but whose relevance is highlighted within these complex morphological systems. These are the problems of Neutralization of categories, Placing of affixes and Licensing of affixes. After a discussion of each of these in general terms, section 3.2 presents a thorough analysis of the Nunggubuyu transitive clitic sequences, with attention to Neutralization (3.2.2) and Placing (3.2.3 - 3.2.4). Neutralization of categories when both arguments are [+participant] is discussed in section 3.3, followed by a brief exposition of the Ket verb-word, with attention to the Placing of the INFL constituent within the complex verb (section 3.4).

3.0.1 The Neutralization Problem

First, if a language has many person-number-gender categories for intransitive agreement, it is usually the case that there are considerable neutralizations in the paradigm for agreement with two or more arguments. At first glance, it might be supposed that that system would be simplest which fully cross-classified all possible intransitive arguments to produce a fully differentiated matrix of agreement for two arguments. In fact, massive neutralizations usually occur, such that, for example, specialized combinations like '1st person exclusive dual acts on 2nd person dual' are rarely if ever encountered, even in languages which have duals for intransitive agreement. I shall refer to this problem as The Neutralization Problem.

Attention was first called to the Neutralization problem for Warlpiri in Hale (1973:330-331). Hale discovered that specifically dual forms exist for all three persons for intransitive agreement in Warlpiri, but the dual tends to be 'replaced' by the plural in certain combinations when both subject and object are nonsingular. The two dialects of Warlpiri, Eastern and Western, differ in regard to how extensive this replacement is. In the Eastern dialect, "if both the subject and object are nonsingular, only plural clitics are allowed":

(1) ngajarra-rlu ka-rna-lu-nyarra nyumpala nya-nyi
we-ERG present-1-pl-2pl you see-nonpast

'We two see you two.'

In Western Warlpiri, however, the facts are more complicated. Dual is replaced by plural obligatorily when both subject and object are dual, and only one dual must be replaced. The dual which is replaced is that which is lowest is the hierarchy 1>2>3:

(2) nyumpala-rlu ka-nku-lu-jarrangku ŋajarra nya-nyi
you-ERG present-2-pl-1du us see-nonpast

'You two saw us two'

A hierarchy of features here dictates the direction of neutralization. Therefore, it seems plausible to assume that the neutralizations in Warlpiri and in other multiple-argument systems results from Impoverishment of morphosyntactic representations which have become too complex, depending on language-particular (or even dialect-particular) constraints on complexity. As discussed in the Introduction, Impoverishment in principle obeys the hierarchy of features and is activated when M^0s are above the complexity threshold for the language.

I will show that not merely do extensive neutralizations occur, but also that they tend to occur in the same places within a paradigm in both Australian and Kiowa-Tanoan, clearly unrelated language families. To explain these facts, I develop a theory of complexity for the morphosyntactic representations of multiple-agreement, whereby neutralizations of this sort can be subsumed under the mechanism of Impoverishment as discussed in the Introduction.

3.0.2 Placing

Second, in those multiple argument agreement systems which are not utterly transparent, that is to say, in which there is not an obvious one-to-one correspondence between syntactic agreement projections and affixes, the distribution of affixes is often described in terms of a template whose slots are keyed to the types of features whose exponents will appear in them. For example, consider the following forms from Western Desert (Australian, Dixon 1980: 362), and Dakota (Schwartz 1979):

(3) a. pu-ngku-rna-nta
hit-FUT-1sgSUB-2sgOBJ 'I will hit you.'

b. pu-ngku-rni-n
hit-FUT-1sgOBJ-2sgSUB 'You will hit me.'

(4) a. ũ-ni-kte
1pl-2PAT-kill 'We killed you (sg).'

b. ũ-ya-kte
1pl-2AG-kill 'You (sg) killed us.'

c. wičʰa-ũ-kte
3plPAT-1pl-kill 'We killed them'

d. wičʰa-ya-kte
3plPAT-2AG-kill 'You killed them'

In both of the above examples, the affixes which realize 1st person features precede those which realize 2nd person features, *regardless* of the syntactic source of those features. In both examples, 1st person affixes (either subject or object) precede 2nd person affixes. In Western Desert, the ordering extends to 1>2>3, but in Dakota, as can be seen in (4c,d), an affix realizing a 3 pl patient precedes both 1 and 2, giving the ordering 3 >2 >1. Considerably more complicated and stipulative such templates have been adduced for the Catalan clitics in Bonet (1991). I will refer to this as the Placing Problem.

Some parts of the Placing Problem can be reduced to phonological considerations, but not all. For example, the fact that the 2sgSUB -n in Western Desert must follow the 1sgOBJ affix -rni in Western Desert is a phonological property of -n. This affix is strictly a suffix within the clitic cluster and therefore cannot be the first member of the agreement clitic sequence. The following examples from Warlmanpa, a related language, show that -n follows the reflexive affix -ñanu, while other agreement markers, both person and number, precede -ñanu (Ken Hale, field notes):

(5) a. =rna-nyanu
1 - refl

b. =lu-nyanu
pl - refl

c. =pala-nyanu
dual - refl

d. =nyanu-n
refl.- 2

In Warlpiri, closely related to Warlmanpa, the 2SUB affix has not reduced to -n but remains fully syllabic as -npa (Hale 1973). The 2SUB affix does not metathesize to the right of a reflexive or of a 1OBJ clitic:

(6) a. =npa-nyanu
2SUB - refl.

b. =npa-ju
2SUB-1OBJ

In the case where metathesis does occur, it is sufficient to assume that the normal Placing order is [subject person > other affixes], but that -n must appear as a suffix to the constituent it is adjacent to within the AGR string:

(7) [-n * [$_Y$ X] --> [X+n]

In the above derivation, -n trades the relation of left-adjacency to the constituent Y (representing a following object or reflexive affix), for the merger relation, whereby it is suffixed to X. We may speak of this merger as being phonologically motivated, on the assumption that no epenthesis rule or other strategy of repair could allow the single segment /n/ to appear in the string as originally positioned.

The Placement of affixes may therefore depend upon their inherent phonological properties, namely, whether they are inherently suffixes or prefixes. The prefix/suffix polarity of an affix is a phonological property of the affix because it is a property relevant to the linearized string phonological information which feeds into PF. Recall that I proposed in chapter 1 that polarity is encoded by means of a feature [±prefix]. I now show that this feature may be underspecified, having its value supplied by the stem to which an affix attaches.

First, affixes may lack the suffix/prefix property and be placed in virtue of the phonological properties of the stem to which they attach.

Fulmer (1991) has shown that certain verbal affixes in Afar, an East Cushitic language spoken in Ethiopia and Djibouti, attach either as prefixes or suffixes. These affixes are prefixal when the verb begins with a nonlow vowel, elsewhere suffixal:

(8) a. t-okm-è
2-eat-perf 'You ate.'

b. yab-t-à
speak-2-impf 'You speak.'

c. ab-t-é
do-2-perf 'You did.'

Afar is unusual inasmuch as the phonological properties of the *stem* and not of the affix determine the Placing of the affix. We may call such affixes *mobile* in the sense that their position is determined for them by other facts about the word in which they appear.

Furthermore, Placing can also be determined by derivational rules. I showed in the Introduction that Huave verbs (and most nouns) require a Theme vowel as part of their autonomous morphological structure. This theme vowel is usually prefixal, but a certain class of verbs have a suffixal theme vowel when reflexive (Stairs &Hollenbach 1981). In the following pairs, a transitive verb meaning 'raise' alternates with a derived intransitive meaning 'rise':

(9) a. a-wit
TH-raise 'He/she raises (it) up'

b. wit-ɨ-m
raise-TH-nonpast 'He/she rises up'

c. t-a-wit
past-TH-raise 'He/she raised (it) up'

d. wit-ɨ-t
raise-TH-past 'He/she rose up.'

e. ap-m-a-wit
fut.aux-fut-TH-raise 'He/she will raise (it) up'

f. ap-wit-ɨ-m
fut.aux-raise-TH-nonpast 'He/she will rise up'

In the transitive forms, (9a,c,e), the theme is in its usual prefixal position; when the verbs are intransitive, the theme is suffixal (9b,d,f). Only derived intransitive reflexives have the suffixal theme vowel: regular intransitives such as a-hɨy 'he/she walks,' or a-kɨik 'he/she laughs' have a prefixal theme. Thus, I postulate a derivational rule of Huave morphology which converts transitives to intransitives as below:

(10) Huave Detransitivization (Lexical)

$$\begin{array}{ccccc} & [+\text{prefix}] & & & [-\text{prefix}] \\ & | & & & | \\ [[\,\text{V}\,] & \text{TH}\,] \longrightarrow & [[\text{V} & + & \text{TH} \\ | & & | & & \\ <\text{Ag}, \text{Th}> & & <\text{Ag}_i, \text{Th}_i> & & \end{array}$$

This rule is a derivational rule operating over the autonomous word structures of Huave. (As a derivational rule, it is only partly productive: not all verbs permit a derived intransitive via a theme polarity change by this rule).

Mobile affixes of this sort are extremely common in Ket, a Paleosiberian language, which will form the focus of discussion at the end of this chapter.

Not only is the theme a mobile affix in Huave, but two other affixes are also mobile and are Placed according to the location of the theme: these affixes are *m* 'nonpast (refl.)/ future' and *t* 'past.' Since the theme is always a vowel, and canonical stems begin and end in consonants, it suffices to say that these mobile affixes, like those of Afar, are placed phonologically, according to where the peripheral vowel happens to be. (Afar presents a more complicated picture, since the relevant distinction is between low and nonlow vowels).

These affixes show overt shifts in its location in (9c,d) for *t* and (9e,f) for *m*. Both must attach to the theme and so are prefixal when the theme is a prefix and suffixal when the theme is as suffix. Not all affixes are mobile in Huave, however. For example, some AGR affixes are invariably suffixes, regardless of the location of the theme:

(11)

a. wit-ɨɨ-t-os
rise-TH-past-1 (past) 'I rose'

b. t-a-wit-ias
past-TH-rise-1 (past) 'I raised (it)'

In the above two examples, the '1 (past)' suffix, which varies harmonically here between -ias and -os, is strictly suffixal, although the past affix varies in its location. Other AGR affixes behave similarly.

To summarize, a number of factors may come into play in Placing affixes. Most typically, the location of an affix will depend upon the location of the syntactic head and M^0 from which it derives. However, when there is M^0 splitting, specific rules or principles of Placing must situate affixes. These rules may be automatic, deriving from the prefix/suffix status of an affix; they may depend upon the phonology of the stem, as in the mobile affixes of Western Desert, Afar or Huave, or their position may show changes in argument structure via a derivational rule, as in the themes in derived intransitives in Huave. Finally, Placing may be stipulative, as in the ordering of clitics in some Romance languages, in particular Catalan. Moreover, Placing may derive from a combinations of these factors, for example a combination of ordering statements and syntactic nestedness, as I will show in section 3.3 for Nunggubuyu. We return to the topic of Strict Placing in section 3.2

The Placing Problem is especially acute in those circumstances where no matter how highly articulated the syntactic agreement projections become, they still cannot place the affixes correctly. For example, as I argued in Noyer (1991), regardless of which order of adjunction is postulated for SUBJECT-AGREEMENT and OBJECT-AGREEMENT phrases in syntax, where Placing obeys a person hierarchy, as in Dakota or Western Desert, one order will logically be a violation of the syntactic order of adjunction. Clearly, for such strings to be derived, the positions of the affixes cannot depend on principles of syntax alone but rather on a more articulated morphological or phonological mechanism of Placing.

3.0.3 Licensing

Third, templates as traditionally employed in fact do more than Placing as in (3) and (4). Their second notational role is to enforce a disjunctivity among affixes competing for a position. I will show that disjunctivity may arise in two ways: (1) features of the input morphosyntactic representation of an M^0 may be Impoverished by filters so that not all "expected" affixes end up appearing or (2) affixes may compete for obligatory positions in autonomous morphological structure. Featural disjunctivity (1) may occur both in systems without obligatory positions-of-exponence, which I will call Free Licensing systems, and in systems with obligatory positions-of-exponence, which will be called Strict Licensing systems. Positional disjunctivity (2), on the other hand, occurs only Strict Licensing systems. First, I will discuss featural disjunctivity.

3.0.3.1 Impoverishment and Competition for Non-obligatory Positions

The first type of disjunctivity arises when Placing appears to put features of the same *type* into the same position, such that two affixes realizing features of this type are 'competing' for the same slot. Below are provided examples of this from Georgian (Kartvel. Vogt 1971) and Maung (Non-Pama-Nyungan Australian, Capell & Hinch 1970):

(12)	a.	v-xedav 1SUB-see	'I see him'
	b.	(*v-) g-xedav (1SUB)- 2OBJ-see	'I see you (sg.)'
(13)	a.	ŋa-i-margudba [> ŋ-i-] 1AG-3PAT.classI-hit	'I am hitting him.'
	b.	(*ŋa-) gu-n-anjandi (1AG)- 2-ptcl-watch.hypothetical	'(Suppose) I watch you.'
	c.	ŋa-ni-n-margudba 1AG-PL-ptcl-hit	'You (pl.) hit me.'
	d.	ŋa-ni-(*ni)-n-margudba 1AG-PL- (PL) ptcl-hit	'You (pl.) hit us.'

In (12b) both the 1SUB prefix and the 2OBJ prefix cannot cooccur, although the 1SUB prefix does appear in (12a), where the 3rd person object has no overt affix. There is in Georgian only one prefix marking agreement, and, where we might expect two prefixes to appear, only one does appear, a prefix indexing the object.

Similarly in Maung, while in (13) the 1AG affix ŋa- may appear along with an affix showing the class of the patient (here, -i- class I), ŋa- cannot cooccur with an affix showing the *person* of the patient, such as gu- '2.' According to the templatic view, there is only one prefix position for person in the Maung agreement clitic sequence, and, when there is competition for this one slot, an affix discharging patient features will win. Likewise, only one plural affix -ni- may appear in any given clitic sequence: where both arguments are plural (participants), only one plural affix may appear, as in (13d).

I will be arguing that these 'positions' are not part of autonomous morphological structure, i.e. they are not Qs in Halle's sense. The crucial assumption is that only obligatory positions may be referred to by morphological rules. Optional 'positions' such as occur in Maung or Berber are not positions but are artifacts of ordering statements holding over the various affixes which an M^0 may split into.

On the traditional view, once a template position is filled, it can no longer accommodate a further string. Thus, besides positioning affixes according to the features they discharge, implicit within the traditional descriptive notation of a template is the *licensing* of these affixes in those positions. In addition to Placing, then, traditional templates perform Licensing. As I will try to show, these are not necessarily related operations, although the template notation leads one to suppose that they are. In fact, I will argue that Licensing and Placing are strictly separate operations.

A major question in the study of systems such as that in Maung is this: are the disjunctions of particular affixes (such as the disjunction of ŋa- '1' and gu- '2' or of more than one instance of -ni- 'pl') the result of competition for slots or are they neutralizations of the input morphosyntactic representation? This problem becomes particularly acute if there is no way formal way to refer to the slots which these affixes are competing for.

Observe that in the Maung case, both affixes competing for a particular slot realize features of the same *type*. Therefore it is equally plausible to assume that there is no competition among *affixes* per se, but rather that an input representation consisting of too many *features* of like type surpasses the complexity threshold for Maung and therefore must be Impoverished before spell-out takes place.

Some principles must determine which features will win out in such a competition: in the Maung case will it be the '1' of the subject or the '2' of the object? A natural assumption to make is that the principles which govern direction of Impoverishment are the same as those which order morphological rules, as proposed in chapter 1. Specifically, object features will beat out subject features on the assumption that objects are more marked than subjects (in a nominative-accusative system).

Recall from the Introduction that I proposed that Impoverishments are effected by filters in Universal Grammar. Essentially, a child will assume that a filter is operative until forced to suppress the filter by positive evidence. Therefore, languages which permit extensive exponence of multiple agreement properties (e.g. Warlpiri) are highly marked, since the automatic neutralization filters operative in Maung, for example, are suppressed. This leads automatically to the prediction that, other things being equal, whenever an agreement system arises which is very rich in the overt exponence of two or more arguments (e.g.Warlpiri), this system should yield diachronically to more compact and less 'expressive' agreement morphology as in Maung.

3.1.1 Strict and Free Licensing

We have seen that the traditional template notation implicitly licenses affixes in certain positions. Since I deny that this traditional notation is anything more than epiphenomenal, I now describe the sources of licensing as I understand them.

Licensing of affixes comes from three sources. The first source is that each M^0 from syntax licenses an affix position: an obligatory Q. This sort of system is exemplified whenever there is a simple isomorphy between M^0s and affixes.

The second licensing source is that autonomous word-structure may require more than one position, as in Arabic, in which case the M^0 splits, filling the independently required positions. In both these cases, an affix may appear only when independently licensed, either by an M^0 from syntax, or by obligatory positions (Qs) which this M^0 must split into for morphological well-formedness. I will call such systems Strict Licensing systems.

The third case, exemplified by Berber and all languages with 'optional' templatic positions, I will call Free Licensing systems, in which each M^0 splits into indefinitely many affixes and no explicit licensing conditions are required for each such affix. In section 1.18 I proposed a parameter, represented formally as a feature [±Autonomous Licensing], which differentiates M^0s such as the Berber AGR and the Arabic INFL. Berber AGR is [-Autonomous Licensing] -- a Free Licensing system -- in which no position is obligatory and as many affixes appear as there are features which can be discharged. Arabic INFL is [+Autonomous Licensing], and each affix may appear only if the word-template is not filled up.

The only real *positional* disjunctivity occurs in [+Autonomous Licensing] systems. Here, a word becomes "filled-up": the word has only so many positions of exponence permitted by its autonomous conditions of well-formedness, and once these are filled, no more affixes may be added to it, regardless of the types of features the affixes express. Recall from Chapter 1 that in classical Arabic the suffix position shows positional disjunctivity: the default tense/mood suffixes *-u/-a/-Ø* appear only when a more specific agreement affix does not fill the suffixal slot.

Because Arabic INFL is [+Autonomous Licensing], then positions are obligatory. It follows that for each position there must exist an explicit default realization: *y-* for the prefix position and *-u/-Ø/-a* for the suffix position, depending upon the tense/mood feature on the stem. Default affixes in obligatory positions are diagnostic of [+Autonomous Licensing] systems.

In Free Licensing on the other hand, affixes are self-licensing, although their location (their Placing) may be stipulated or automatic, depending upon other factors (discussed in section 3.1.2). I proposed in the last section that apparent disjunctivity of affixes within *non*-obligatory positions results in all cases from Impoverishment and not from 'competition' for slots.

As an example of a Free Licensing system, consider again the Berber conjugation as analyzed in section 1.18. (Although the Berber paradigm below is not a multiple-argument agreement system, it displays Free Licensing). Berber also presents the case where an M^0 may split indefinitely, and hence there are no explicit licensing conditions. The conjugation for the Tamazight dialect is repeated below:

(14) Tamazight Berber Prefix Conjugation *dawa* 'cure'

	sg	pl
3m	i-dawa	dawa-n
3f	t-dawa	dawa-n-t
2m	t-dawa-d	t-dawa-m
2 f	id.	t-dawa-n-t
1	dawa-ɣ	n-dawa

Tamazight Berber verbs may have up to three affixes, as in t-dawa-n-t '2-cure-pl-f: you (f pl) cure' But no single one of these affix positions is obligatory (in the sense that each position need not have an overt filler): verbs may have no prefix and only one suffix, as in dawa-ɣ 'cure-1: I cure' or a prefix and no suffix: i-dawa 'sg.m-cure: he cures.' Because Berber AGR is [-Autonomous Licensing], no position-of-exponence is obligatory. This means, in effect, that if Berber AGR were *not* 'liberated' in this way, it would require three independently licensed positions, each with a Ø elsewhere affix. Such zeroes, however, are unmotivated and have no content. Therefore, Berber AGR is liberated and splits as much as it can.

In a Free Licensing system, there are as many positions of exponence as there are rules which discharge features as affixes. Each affix licenses its own position, and there is therefore never any competition for slots.

For example, consider the Berber form dawa-n-t 'cure-pl-f: they (f) cure.' Because the plural affix *-n* is NOT disjunctive with the feminine affix *-t*, we may say that both affixes are licensed. But this licensing in not because there are two suffix positions-of-exponence in Berber, in the sense of positions-of-exponence as *obligatory* elements of autonomous word-structure (as in Arabic). Rather, Berber has no explicit licensing conditions at all: each affix licenses itself. Because there are no slots for which affixes are in competition in a Free Licensing system, there can be no *position-based* disjunction among affixes, although, as I showed in section 1.18, discontinuous bleeding continues to apply, such that features are discharged only once. Spell-out rules thus continue to apply until as many of the features of the input are discharged as can be.

In Maung as well, prefix positions are not obligatory. In Maung, if there is no 1st or 2nd person argument, the person prefix positions may be empty (15a); if no participant is plural, then the position for the affix -ni- may also be empty (15b):

(15) a. Ø-Ø-ji-wu-ŋar̲undin
3classI-3classIII.ERG-cook in hole.PAST 'They roasted it.'

b. gu-Ø-n-Ø-ŋa-wa̠lgu-ŋ
2-ptcl-3classII.ERG-promise-PAST 'She promised you.'

In such cases, the 'positions' do not contain contentful Øs. These empty Øs represent only places in the string where other affixes *would have* appeared if the AGR had different features.

Iin Maung, one prefix 'position' is specifically *keyed* to person features and one to the feature [-sg]: only affixes which are primary exponents of (at least one) of these features may appear in these positions. On the other hand, the Arabic suffix position is neutral with respect to the type of features which may appear in it: it does not matter, as long as the position is eventually filled. In Berber, finally, 'positions' are neither keyed nor obligatory.

Therefore, we could write a template for Maung, for example, that has an optional Person-Prefix slot and an optional Number-Prefix slot and two optional Class prefix slots. For Berber, we could write a 'template' with three unobligatory slots. For Arabic, on the other hand, there is an obligatory, unlabelled prefix and suffix slot:

(16)	Maung:	(Q_{Person}) (Q_{Number}) (Q_{Class}) (Q_{Class}) Verb
	Berber:	(Q) Verb (Q) (Q)
	Arabic	Q Verb Q

I proposed in chapter 1 that the Arabic system is derived by conditions on the well-formedness of words, and that these conditions are encoded in an autonomous module of grammar figuring in the mapping from syntax to phonology. This is one way of defining a template: a mold which must be *filled up* in order to give a well-formed object.

The Maung and Berber cases are different: since the Qs are not obligatory, it makes no sense to say that the *positions* are required for morphological well-formedness. The notational function of the Maung 'template' is merely to order the affixes appropriately and to ensure that there is only one affix of each type. This represents another conception of a template: a mold which enforces a particular form on an object.

I propose that these ordering facts are to be accounted for by separate principle of Placing which I discuss in the next section; I have already suggested that Impoverishment ensures that only one affix of each type appears in the Maung prefix string. Therefore, the template has no independent function but is a theoretical artifact in a Free Licensing system.

To summarize, I now present the following hypothesis:

(17) M^0s have the feature [±Autonomous Licensing]

If M^0 = [+Autonomous Licensing], then affixes discharging features in M^0 appear only in obligatory positions (Qs) appearing in autonomous morphological structure. Each Q may be realized as only one affix; therefore there may be positional disjunctivity

If M^0 = [-Autonomous Licensing], M^0 may split into as many affixes as there are relevant rules to discharge the features of the M^0. Apparent positional disjunctivity arises solely from Impoverishment.

The differences between Maung, Arabic, and Berber are now shown according to the characteristics below:

(18)

	[±Aut. Lic.]	Placing	Qs and default affixes	neutralization of like features
Arabic	+	-	+	-
Maung	-	+	-	+
Berber	-	-	-	-

(19) [+Aut. Lic.] --> obligatory positions, default affixes, competition for slots

[-Aut. Lic.] --> unlimited splitting, optional 'positions'

Turning now again to the examples in section 3.0, it is now apparent that Western Desert and Dakota, like Berber, are Free Licensing systems (at least with regard to verbal agreement). In both Dakota and Western Desert, there is no competition for slots and no position is obligatory.

In terms of acquisition, the Strict Licensing system is discoverable whenever there exist non-zero elsewhere rules, that is, affixes whose content is not a natural class. Strict Licensing is also discoverable when there is competition for *obligatory* slots among independently occurring affixes, such as the Arabic person-number-gender endings *-na, -aani*, *-iina*, *-uuna* disjunctive with the imperfect ending *-u*. Free Licensing is discoverable whenever there is manifest splitting of M^0s into non-obligatory positions, as in Berber or Maung.

To know Arabic, one must know that each nonperfect verb has at most and at least one prefix and suffix. This fact must be learned as part of the autonomous morphological structure of a certain type of Arabic INFL.

To know Berber, one must know that the AGR M^0 splits variously and indefinitely, and that the expected one-to-one relationship between M^0s and affixes is waived. This too must be learned as the [-Autonomous Licensing] parameter of Berber AGR. Beyond this parameter, one need know for Berber only the verbal agreement affixes and the rule hierarchy which dictates their order of attachment and relative rank.

To know Maung, one must know that the AGR M^0 splits into a number of non-obligatory positions and is therefore also [-Autonomous Licensing]. Additionally, however, one must know both the principles which correctly Place affixes and the Impoverishments which prevent more than one person or number affix from being generated.[1] We now turn to the mechanisms of Placing.

3.2 Strict or Free Placing

Placing may also be Free or Strict. With Free Placing, affixes attach in the order predicted by the hierarchy of features which governs morphological rule ordering. With Strict Placing, language-particular stipulations govern the positions of affixes.

In chapter 1, I showed that Arabic is a Free Placing system, but within a Strict Licensing system for positions of exponence. There are only two positions of exponence for the prefix-conjugation: the prefix, and the suffix. The prefix/suffix property of each affix dictates its position in the string. There is no need for specific rules of Placing. Insofar as the rules of Placing would have to mirror the prefix/suffix information in each affix, there would be a considerable duplication if separate placing rules were to position these affixes' features and then to realize them as strings.

Next, consider again the Berber form t-dawa-n-t '2-cure-pl-f: they (f) cure.' The two suffixes appear in the order pl > f, just as is predicted by the hierarchy. The rule discharging 'pl' applies first and places the suffix at the right edge of the stem. The rule discharging 'f' then applies placing *-t* to the right edge of this already suffixed form:

[1] Put more correctly, one must know which Impoverishment filters are not suppressed: most filters come for free and are 'unlearned.' As I proposed in the Introduction, filters are assumed automatically by the child unless positive evidence is provided to switch them off.

(20)	Rule	Representation	Morphological Word
	Input	[[V] 2 pl f]	
	2 t-	[[V] pl f]	t -V
	pl -n	[[V] f]	t -V-n
	f -t	[[V]]	t -V-n-t

What is crucially the case is that the Placing and the spelling-out/discharge of features happen *simultaneously*. All Placing in a Free Placing system can be derived automatically from the order of the rules and the suffix/prefix property of the affixes.

Thus, in a Free Placing and Free Licensing system, the derivation resembles most closely the Extended Word and Paradigm model of Anderson (1981) and subsequent work. The difference between that model and the theory advanced here is that the rules in (20) actually discharge features of the input, so as to allow bleeding without rule block disjunctivity (recall section 1.18.3).

The hierarchy of features determines the ordering of affixes in a Free Placing system. Therefore, we should expect that there is some correlation between the level of embedding of affixes in a Free Placing/Free Licensing system (such as Berber) and the 'winner' affixes when affixes compete for positions in a Strict Licensing system such as Arabic. This can be diagrammed as below:

(21)				
	a.	.... p] g] n]		Free Placing/Free Licensing
	b.	x	'p'	Strict Licensing
		y	'g'	
		z	'n'	

In a Free Placing/Free Licensing system, affixes realizing person (p), number (n), and gender (g) should have the (unmarked) level of embedding as in (21a), since the person affix will attach first, and then the number affix, etc. If all three affixes are competing for the same slot in a Strict Licensing system, the the 'p' affix should win, since the rule realizing this affix will apply first. Thus, the Feature Hierarchy Hypothesis can now be amended:

(22) *The Feature Hierarchy Hypothesis*

There is a universal hierarchy of morphosyntactic features.
If F and G are morphosyntactic features and F is higher than G on the hierarchy, then:

(1) if *[αF βG] is active at Morphology, then [αF βG] is Impoverished to [αF]

(2) if two spell-out rules, one referring to F, the other to G and not to F, have disjoint or overlapping structural descriptions, then the rule referring to F applies first

Corollary of (2): in a Free Licensing/Free Placing system, an affix realizing F will appear more embedded than an affix realizing G but not F

The corollary above has an interesting parallel in the observation of Bybee (1985:13):

(23) These results [i.e. of a study of morpheme ordering, R.N.] suggest a "diagrammatic" relation between meanings and their expression, such that the "closer" (more relevant) the meaning of the inflectional morpheme is to the meaning of the verb, the closer its expression unit will occur to the verb stem. This type of diagrammatic relation is also evident in the degree of fusion between the expression of the verb stem and the inflectional morphemes....

For Bybee's conception of 'relevance' I am substituting the notion of feature hierarchy. In general, the higher on the hierarchy the features are which an affix realizes, the closer to the stem (including fusing with the stem) it will occur in the unmarked case, i.e. a Free Placing system. The Feature Hierarchy Hypothesis is more predictive than Bybee's observation because it correlates neutralization, affix order, and the 'winners' in competition for strictly licensed slots.

Turning now to Strict Placing, on the other hand, this occurs when specific ordering statements (traditionally notated by means of templates) situate affixes in a string. An example of Strict Placing can be found in the positioning of clitics in Catalan. Bonet (1991:103) proposes the template in (24) for the Barceloní dialect of Catalan. Mapping to this template linearizes clitics according to their feature content. Position 1 is occupied by the reflexive (no marked person features), position 2 by the 2nd person, position 3 by the first person, position 4 by 3rd person (unmarked), position 5 by the genitive (partitive) and position 6 by the neuter or oblique. Sentence (25) shows all positions with an overt clitic:

(24)

CL	CL	CL	[ARG]	[GEN]	[OBL]
					or
ARG	ARG	ARG			[NEUT]
[PERS]	[PERS]	[PERS]			
	[-1]	[+1]			

(25) [sə tə mə lzə n i] *vas* *quedar tres*
refl-2-1-3pl-gen-obl aux-2sg take three

'You took three of them from mine (e.g. children).'

The ordering of the clitics in (24) is largely stipulative. Bonet considers but rejects syntactic solutions to the placing problem. A decisive argument showing how stipulative the ordering of clitics is in Catalan comes from Bonet's examination of dialectal variation:

> While in Barceloní the order of between first and second person, and the bare {PERSON] clitic *es* is as ... for Standard Catalan (es-II-I), the order in Gironí, just a few miles to the north) is II-I-*es*. In addition, in dialects very close to Barceloní ... the order is II-*es*-I. Nothing in the syntax of these dialects (identical as far as I know) could account for this difference (Bonet 1991:75).

To Bonet's assertion that syntactic variation cannot account for this difference in clitic ordering we may add as well that the hierarchy of features cannot be adjusted to predict all three different orderings II-I-*es*, II-*es*-I and *es*-II-I. At least two of them must be stipulated orderings, if not all three.

How shall such stipulated orderings be effected? We can construct two hypotheses regarding how Strict Placing occurs. These will be called the Representational view and the Derivational view. Bonet's template for Catalan exemplifies the Representational view. Speakers of Barceloní will 'know' the representation in (24) and the principles which dictate how morphosyntactically represented clitics link to this template. Thus, in virtue of knowing the representation of the template and clitics, the placing of the clitics is derived by mapping principles.

On the Derivational view, the templatic representation in (24) is epiphenomenal. Instead of knowing the representation of the template, speakers instead know a set of ordering relations among the different clitics (or, equivalently, among the rules which realize the clitics). For example, the template in (24) can be generated from principles of ordering, which will include statements like "an [ARG] clitic will precede any non-[ARG] clitic; any [PERSON] clitic will precede any non-[PERSON] clitic, etc." The Catalan case is particularly specialized, so that the number of discrete statements required to uniquely generate this template is large. However, this is an advantage of the analysis, since the Catalan case is quite exceptional.

A further advantage of the Derivational view of template 'construction' is that certain of the discrete statements involved in generating a template can be evaluated as less or more costly than others. For example, the Maung template can be derived according to the statements: person > number, number > class, as is not atypical for Australian clitic clusters; similarly the Dakota template is derivable from the statement 1>2. These discrete ordering statements are unmarked (and therefore costless) options of Universal Grammar; other discrete ordering statements, such as perhaps the 2>1 ordering of Barceloní Catalan, as opposed to the 1>2 ordering in Valencian (Bonet 1991:131), are marked and therefore costly. Any given template will combine marked and unmarked options among its discrete ordering statements. Furthermore, if templates are generated by sets of discrete ordering statements, then acquisition and historical change will add or modify these statements individually to a grammar.

On the Representational view, there is no obvious way to decompose a template and therefore to evaluate its cost or describe its historical or psychological development.

Free Placing can also be understood within a Representational or Derivational View. On the Representational view, Free Placing systems merely have no templates: they are 'auto-templatic.' On the Derivational view, Free Placing entails, in effect, that each affix automatically creates its own discrete placing statement. For example, the Berber rule/suffix *-n* 'pl', in virtue of its being a *suffix*, automatically generates a placing statement which places the feature 'pl' to the right of the existing form and then discharges this feature and spells it as *n*. What is special about these automatically generated placing statements is that they situate the affix only vis-à-vis the pre-existent stem and not vis-à-vis any other *particular* affix. Thus, Free Placing is derived only from the suffix/prefix property of an affix and not in virtue of any other ordering statement. So we can call Free Placing systems 'auto-regulatory' as well.

The order of the rules introducing affixes is the same as the sequential order these affixes come to have in a template (or the level of embedding if systems with prefixes and suffixes). Therefore, language-particular discrete ordering statements such as Catalan 2 > 1 are in fact extrinsic ordering statements (costly, learned) for the rules which introduce the 2 affix /t/ and the 1 affix /m/. On this view, which I now adopt, 'templates' arise from the order in which morphological rules apply; this order may be entirely unmarked, obeying the hierarchy of features, in which case we have a Free Placing system, or it may have extrinsic and stipulative ordering, in which case we speak of Strict Placing.

To summarize, the ordering of clitics in Catalan represents a system of Strict Placing, whereas, insofar as the order of affixes obeys only the hierarchy of features, the Berber system represents Free Placing, and, insofar as the prefix/suffix property of the affixes determines their location, the Arabic system represents Free Placing as well.

One very desirable result of the Derivational view of template construction is that it allows the decomposition of a template into the discrete ordering statements that generate it. This then allows the ordering statements to have different sources: stipulative, phonological, or syntactic. In the previous section, I reviewed the phonological motivations for affix Placement in Afar, Western Desert, Warlmanpa, and Huave. We now return to the Dakota ordering 3 > 2 > 1 introduced earlier and see that this ordering derives from the interaction of two discrete statements of differing source.

Schwartz (1979) identifies two competing analyses for the observed orderings in Dakota, which are summarized in the table below, taken from Schwartz (1979:8):

(26)		3rd person >	1st person >	2nd person	
	a.	wič^ha	wa		1: 3pl
	b.	wič^ha	ũk		1pl : 3pl
	c.	wič^ha		ya	2 : 3pl
	d.		ma	ya	2 : 1
	e.		ũk	ya	2 : 1pl
	f.		ũk	ni	1pl : 2

On the first analysis, the positions of affixes is determined by the ordering relation 3 > 1 > 2. Thus, 3 > 2 in (26c), 3 >1 in (26a,b), and 1 > 2 in (26d,e,f). However, because the only 3rd person clitic that exists in the transitive agreement systems is the 3pl *patient* clitic wič^ha, it is also the case that, with only one exception, patient clitics precede agents clitics. The one exception is when the agent is 1st person and the patient is 2nd person (26f). Rather than state the ordering generalization as 'patients precede agents' and include an exception statement, Schwartz argues that the affixes are positioned according to 3 > 1 > 2, which she takes ultimately to be a condition on the individual affixes.

However, there is another factor at play here: observe that the 1pl affix ũk is used when indexing *either* the patient and agent. It is therefore unmarked with regard to its syntactic source. Thus, the ordering in (26f) is not in fact at variance with the principle that patient affixes precede agent affixes, whenever these affixes specifically index the patient or the agent. This observation allows the following straightforward account of the Dakota ordering to emerge.

I proposed at the beginning of the section that, in the least costly instance, the order of application of morphological rules is a function of the hierarchy of features. Some hierarchy, whether marked or unmarked, will by hypothesis derive the ordering in Dakota. Two ordering principles derive the correct results:

(27) a. patient > agent
b. [+I] > [+you]

The independence of these two principles can be observed in 'double-patient' constructions of Lakhota, where neither argument of a transitive clause "performs/effects/instigates or controls" the action (Mithun 1991:517). In the examples below, taken from (Mithun 1991), the 1st person patient *ni* and 2nd person patient *ma* cooccur in that order. Under such circumstances, principle (27a) cannot order the clitics at all:

(28)	a.	iyé-ni-ma-č^{h}eča	'I look like you'
	b.	í-ni-ma-skokeča	'I am as large as you.'
	c.	iyó-ni-ma-kip^{h}i	'I find you congenial'

Separating the two principles then allows them to be evaluated individually in terms of cost. As for (27a), we return to the matter of agent/patient marking in chapter 4. Considering for now only (27b), this ordering is either costless (if the clitic cluster is spelled left-to-right as a separate M^0) or costly (in the clitics are added form right-to-left as affixes to the verb). A more thorough investigation of the morphology of Dakota would be necessary to determine which scenario is correct. However, what is clear is that the Dakota 'template' 3 > 1 > 2 is an epiphenomenon arising from the interaction of the two principles in (26). I will show in section 3.2.3 that the same sort of interaction among Placing principles occurs in Nunggubuyu and motivates in large part the extensive neutralizations seen in its agreement sequences.

Schwartz's conclusion that the individual affixes (their *form* rather than *function*) determine their placement is partly correct. It is the featural content of each affix which determines its placement, regardless of its *de facto* function in the system. In particular, the feature content of ũk is merely '1 pl' without patient or agent function. Therefore this affix may precede a patient affix as in (26f), although it discharges features associated with the agent. On the theory adopted here this is not surprising: the rules which discharge features as affixes refer to as few features as necessary to determine the proper realization of an affix.

To summarize, the following parameters classify the systems so far observed:

(29) Typology of M^0 Splitting Systems

	neutralization of like features	autonomous obligatory positions	costly discrete ordering statements
Berber	-	-	-
Arabic	-	+	-
Maung	+	-	-
Dakota [2]	+	-	?
Catalan	+	-	+

We have already observed systems of the first two types-- Arabic and Berber -- and the last type, Catalan, as exemplified extensively in Bonet (1991). The remainder of the chapter addresses those systems like Maung and Dakota that fall in the gray zone, where there is neutralization of like features, and/or affixes are positioned by unmarked principles of hierarchy or by marked discrete ordering statements, or a combination of such principles. The first such system to be discussed is Nunggubuyu.

[2]Pluralization of 2nd person (agent or patient) and 3rd agents occurs via suffixation of *-pi* in Dakota. When 3rd plural acts on 2nd plural patient, we might expect to find two instances of *-pi*. However, this does not happen. The following forms, taken form Buechel (1970:47), shown that one instance of *-pi* occurs even when both arguments are plural:

(i) ni-cǎštaka-pi 'they/he strike(s) you (sg,pl)'
2-strike-pl

(ii) ni-yátaN-pi 'they/he praise(s) you (sg,pl)'
2=praise-pl

The above forms realize both singular or plural subjects or objects when 3 acts on 2. This suggests that there may be neutralization of like features within positions in Dakota, which therefor may present a mixed system, partly like Arabic and partly like Maung.

3.2 Transitive Agreement in Nunggubuyu

In this section, I propose an analysis of the agreement clitic prefix cluster in Nunggubuyu, a non-Pama-Nyungan language of Arnhem Land, Northern Territory, Australia. All data in this section come from the work of J. Heath (1980, 1982, 1984), who has published an extensive grammar, dictionary, and texts of the language based on his own field work. Textual examples from Heath (1980) are given by section and line following the sign '§'.

From the point of view of agreement systems, Nunggubuyu is remarkable, even within Australia, for its complexity. There are no less than 22 person-number-gender categories which participate in agreement, resulting from a full array of four person categories, [±aug] number, as well as three human genders and six noun classes. These genders and classes present a complexity not found in the Pama-Nyungan languages of Australia discussed above, such as Warlpiri or Western Desert.

Logically, there exist 484 potential combinations of two arguments in the transitive clitic cluster, of which 112 are semantically reflexive (e.g. 'I acting on me') and therefore do not arise. This leaves, however, 372 potential combinations, by no means all of which actually occur as specific forms. The actual number of generable forms is around 130 or so, depending on the analysis. (Each of these also has a variant occurring in particular tense/mood categories which Heath calls the B forms of the clitic. I will delay consideration of the B forms until section 3.2.4, after the basic, A forms are analyzed). Because of the extraordinary richness of the system and the extensive neutralizations within it, Nunggubuyu provides an excellent source of data on the Neutralization Problem announced in section 3.0.1.

I will show that these neutralizations can be expressed as Impoverishments operating at the featural level. Some of these Impoverishments are motivated system-internally to ensure that no strings of affixes will be generated that cannot come to be ordered correctly by the various ordering requirements of the affixes. This is a particularly interesting result inasmuch as it shows that operations within the split M^0 are complex and highly rule-governed.

3.2.1 Overview of Categories and Intransitive Agreement

In section 2.2.6 I introduced the agreement system of Nunggubuyu, with attention to explaining how a trial interpretation could arise by functional inference in the 1st inclusive, and a dual in the same manner in other categories. From the data in Heath

(1984), I cited the following twelve [+participant] arguments. In the far right column I now give the abbreviation that will designate these categories in the following discussion:[3]

(30) Participant Categories in Nunggubuyu

category	intr. prefix	[± I]	[± you]	[± aug]	PLUR	MASC	FEM	abbrev.
1 sg	ŋa-	+	-	-				1-
1 excl m dual	nV: - ni-	+	-	+		m		1+m
1 excl f dual	nV: - ŋi-	+	-	+			m	1+f
1 excl plur	nV: - ru-	+	-	+	m			1+pl
1 incl dual	na-	+	+	-				12
1 incl m trial	ŋV: - ni-	+	+	+		m		12+f
1 incl f trial	ŋV: - ŋi-	+	+	+			m	12+m
1 incl plur	ŋV: - ru-	+	+	+	m			12+pl
2 sg	nun-/ba-	-	+	-				2-
2 m dual	na-wV-ni-	-	+	+		m		2+m
2 f dual	na-wV-ŋi-	-	+	+			m	2+f
2 plur	na-wV-ru-	-	+	+	m			2+pl

All the categories abbreviated with a [+] are [+augmented] categories, those labelled [-] are [-aug] categories; 1 = [+I], 2 = [+you], and m, f, and pl are the privative class attributes. These symbols will provide a convenient index of the feature composition of each category or affix.

Nunggubuyu also has ten [-participant] or 3rd person arguments. These can be arranged according to the eight noun genders of the language, which Heath labels masculine singular, feminine singular, plural, NA, NgARA, WARA, ANA and MANA. The masculine and feminine genders also have [+aug] counterparts which are functionally 'dual' categories.

The masculine (singular), feminine (singular) and plural genders refer to humans. Their nonhuman noun class counterparts are NA, NgARA, and WARA respectively. By this it is meant that, for the purposes of verbal agreement, in nearly all cases, masculine = NA, feminine = NgARA and plural = WARA. Therefore, it is possible to treat these genders as [-human] counterparts of the [+human] genders MASC, FEM, and PLUR (recall that *plural is*

[3]Note again that Nunggubuyu has two sorts of surface [w]. One alternates with [g] and will be written [ɰ], the other alternates with [b] and will be written as plain [w]. These two correspond to Heath's [w_1] and [w_2] respectively. I am avoiding his subscript system because it is no longer typographically necessary. There is no surface difference between the [w]s derived from these segments.
In terms of the feature geometry model of Sagey (1986), modified in Halle (1989), w_1 is dominated by the DORSAL node and w_2 by the LABIAL node. A similar series of glides and alternants occurs in Fula.

a gender value in Nunggubuyu: this may be confusing). Nonhuman genders do not have functional duals. The chart below displays the classes and their intransitive prefixes, along with the abbreviation I shall employ on the far right; [*] means nonhuman, and is usually not relevant to verbal agreement.

Observe that the plural gender categories are automatically [+augmented]; this is notated by a [+] in their abbreviation. The two gender attributes ANA and MANA will be called CLASS features, whereas MASC, FEM, and PL, which may apply to humans or nonhumans, will be called GENDERs. Strictly speaking, both GENDER and CLASS are genders features in the sense of Aronoff (1992), since both propagate agreement in syntax. The genders which I will be calling CLASS are really those genders which are incompatible with the value [+human].

(31) Eight Noun Genders in Nunggubuyu

	prefix	[±human]	MASC	FEM	PLUR	ANA	MANA	abbrev.
masc sg	ni	+	m					m
fem sg	ŋi	+		m				f
plur	wV-ru	+			m			+pl
NA	ni	-	m					m*
N^{g}ARA	ŋi	-		m				f*
WARA	wV-ru	-			m			+pl*
ANA	ɰu	-				m		A*
MANA	ma	-					m	M*

In addition, the singular human categories also permit [+aug] variants, which, like the plural gender categories, have the [+aug] affix wV-:

	prefix	[±human]	[±aug]	MASC	FEM	PLUR	abbrev.
masc dual	wV-ni	+	+	m			+m
fem dual	wV-ŋi	+	+		m		+f

We can now identify the following affixes within agreement system, following Heath's parsing of the forms:

(32) Affixes of the Nunggubuyu Prefix Agreement System

	Underlying form	Surface Instantiations
a. 12+	ŋV:	ŋi: , ŋu: , ŋa :, ŋi , ŋu , ŋa
b. 12	nV	na, ni
c. 1+	nV:	ni: , nu: , na :, ni , nu , na
d. 1	ŋa	ŋa, ŋu
e. 2+	na	na, nu, ni
f. 2-	nun / ba	nun, num, nuŋ, nu, ni, ba
g. +	wV	wa, wu, wi, a, u, i, Ø, ba, bu, bi
h. m	ni	ni, nu
i. f	ŋi	ŋi, ŋu
j. pl	ru	ri, ru, ra
k. A	ɰu	wu, gu, Ø[4]
l. M	ma	ma

These affixes comprise the building blocks out of which the transitive agreement sequences are formed. When appearing as transitive agent or patient affixes, some show a different vowel quality or are slightly truncated, but their approximate shape is always a variant of the underlying shape shown at left. The augmented affix -wV- shows the most variation. Its initial /w/ hardens to /b/ after a nasal but disappears intervocalically. When following a long vowel, the remaining -V- also deletes, so sometimes this affix is utterly deleted by regular phonological rules. Deleting segments are notated in square brackets.

Before proceeding to the transitive clitic sequences, it will be useful to consider the cooccurrence restrictions on the features which comprise the 22 intransitive agreement categories. These restrictions hold from d-structure onwards, that is, no combination of features violating them will compose a well-formed argument in syntax.

First, all [+participant] arguments are [+human], so that, for example, there are no categories *specifically* for addressing or including an entity of the nonhuman genders, such as ANA or MANA:

(33) *[-human +participant]

[4]ANA class appears as Ø when marking the object and following the affix -rV- 'pl'.

Second, the CLASS attributes ANA and MANA do not cooccur with the value [+human], that is to say, ANA and MANA do not have human counterparts within the gender system, whereas NA, NgARA and WARA do:

(34) *[+human CLASS]

Third, the nonaugmented participant categories do not permit gender distinctions. In other words, there is no masculine or feminine 1 sg, 1 incl dual, or 2 sg:

(35) *[+participant -aug GENDER]

Furthermore, all arguments with PLUR gender are automatically [+aug]. We will see that often PLUR gender is a default enhancement of [+aug] and that these features share a certain relationship arising from this.

(36) *[-aug PLUR]

Finally, the GENDER and CLASS features are always mutually exclusive, such that there is no masculine ANA argument or plural MANA argument, for example. To derive this, we may assume that each argument permits at most one such attribute.

Finally, the features align in a hierarchy as below:

(37) [I] > [you] > [augmented] > GENDER > CLASS

This hierarchy is consistent with all hierarchies so far proposed, bu includes an extra distinction at the bottom between those genders which can refer to humans (GENDERs) and those which cannot (CLASSes). This final part of the hierarchy is clearly derived from a further hierarchy of the form [+human] > [-human].

3.2.2 Neutralizations in Transitive Agreement

When subjects and objects combine to form a complex transitive agreement M^0, considerable neutralizations take place. The full set of basic transitive agreement sequences are given in Tables I -IV (pp. 251-255). Some neutralizations are obvious from the table: these are enclosed in boxes. All forms within a box neutralize to the one form, which instantiates the category for the slot in which it sits. Other neutralizations arise from certain

TABLE I: Transitive Agreement, Participant Objects (A forms)

Participant Objects (A forms)

S ↓ O →	1-	1+	12	12+(pl)	2-	2+
1-	-	-	-	-	ŋu-nu 1.2-	ŋa-na 1.2
1+	-	-	-	-		
2-	numba *or* nimba		-	-	-	
2+f			-	-	-	-
2+m			-	-	-	-
2+pl	ŋiri		-	-	-	-
f	ŋa-[N]-ŋi 1.INV.f	na-[N]-ŋi 1+.INV.f	na-[N]-ŋi 12.INV.f	ŋa:-(ra)-[N]-ŋi 12+.(pl).INV.f	ni-[N]-ŋi 2-.INV.f	na-[N]-ŋi 2.INV.f
m	ŋa-[N]-ni 1.INV.m	na-[N]-ni 1+.INV.m	na-[N]-ni 12.INV.m	ŋa:-(ra)-[N]-ni 12+.(pl).INV.m	ni-[N]-ni 2-.INV.m	na-[N]-ni 2.INV.m
+	ŋa-m-bi 1.INV.+	na-m-bi 1+.INV.+	na-m-bi 12.INV.+	ŋa:-(ra)-m-bi 12+.(pl).INV.+	ni-m-bi 2-.INV.+	na-m-bi 2.INV.+
A	ŋa-ŋ-gu 1.INV.A	na-ŋ-gu 1+.INV.A	na-ŋ-gu 12.INV.A	ŋa:-(ra)-ŋ-gu 12+.(pl).INV.A	nu-ŋ-gu 2-.INV.A	na-ŋ-gu 2.INV.A
M	ŋa-[N]-ma 1.INV.M	na-[N]-ma 1+.INV.M	na-[N]-ma 12.INV.M	ŋa:-(ra)-[N]-ma 12+.(pl).INV.M	nu-[N]-m 2-.INV.M	na-[N]-ma 2.INV.M

TABLE II: Transitive Agreement, Nonparticipant Objects (A forms)

S O→ ↓	f	m	m*	+m,+f,+pl +pl	A	M
1	ŋa-ŋu 1.f	ŋa-nu 1.m	ŋa-ni 1.m*	ŋa-[V]-ra 1.+.pl	ŋa-ɰu 1.A	ŋa-ma 1.M
1+f		nu:-ŋu-nu 1+.f.m	ni:-ŋi-ni 1+.f.m*	na-V-ŋi 1+.+.f	ni:-ŋi-wu 1+.f.A	ni:-ŋi-ma 1+.f.M
1+m				na-V-ni 1+.+.m	ni:-ni-wu 1+.m.A	ni:-ni-ma 1+.m.M
1+pl	ni[:]-ri-ŋa 1+.pl.f	nu:-nu 1+.m	ni:-ni 1+.m*	na[:]-[V]-ra 1+.+.pl	ni[:]-ri-Ø 1+.pl.A	ni[:]-ri-ma 1+.pl.M
12	na-ŋu 12.f	na-nu 12.m	ni-ni 12.m*	na-[V]-ra 12.pl	na-ɰu 12.A	na-ma 12.M
12+f		ŋu:-ŋu-nu 12+.f.m	ŋi:-ŋi-ni 12+.f.m*	ŋa[:]-V-ŋi 12+.+.f	ŋi:-ŋi-wu 12+.f.A	ni:-ŋi-ma 12+.f.M
12+m				ŋa[:]-V-ni 12+.+.m	ŋi:-ni-wu 12+.m.A	ni:-ni-ma 12+.m.M
12+pl	ŋi[:]-ri-ŋa 12+.pl.f	ŋu:-nu 12+.m	ŋi:-ni 12+.m*	ŋa[:]-[V]-ra 12+.+.pl	ŋi[:]-ri-Ø 12+.pl.A	ŋi[:]-ri-ma 12+.pl.M
2	nu-ŋu 2-.f	nu-nu 2-.m	ni-ni 2-.m*	ba-ra 2-.pl	nuŋ-gu 2-.A	nu-ma 2-.M
2+f		nu-V-ŋu-nu 2.+.f.m	ni-V-ŋi-ni 2.+.f.m*	na-V-[V]-ŋi 2.+.+.f	ni-V-ŋi-wu 2.+.f.A	ni-V-ŋi-ma 2.+.f.M
2+m				na-V-[V]-ni 2.+.+.m	ni-V-ni-wu 2.+.m.A	ni-V-ni-ma 2.+.m.M
2+pl	ni-[V]-ri-ŋa 2.+.pl.f	nu-V-nu 2.+.m	ni-V-ni 2.+.f	na-[V]-[V]-ra 2.+.+.pl	ni-[V]-ri-Ø 2.+.pl.A	ni-[V]-ri-ma 2.+.pl.M
f	ŋu f	ŋu-nu f.m	ŋi-ni f.m*	wa-ŋi +.f	ŋi-ɰu f.A	ŋi-ma f.M
m		nu m	ni m*	wa-ni +.m	ni-ɰu m.A	ni-ma m.M
+f		wu-ŋu-nu +.f.m	wi-ŋi-ni +.f.m*	wa-V-ŋi +.+.f	wi-ŋi-wu +.f.A	wi-ŋi-ma +.f.M
+m				wa-V-ni +.+.m	wi-ni-wu +.m.A	wi-ni-ma +.m.M
+pl	wi-ri-ŋa +.pl.f	wu-nu +.m	wi-ni +.m*	wa-[V]-ra +.+.pl	wi-ri-Ø +.pl.A	wi-ri-ma +.pl.M
A	ŋi-ŋ-gu f.INV.A	ni-ŋ-gu m.INV.A	Ø	wa-ŋ-gu +.INV.A	Ø	Ø
M	ŋi-[N]-ma f.INV.m	ni-[N]-ma m.INV.M	Ø	wa-[N]-ma +.INV.M	Ø	Ø

TABLE III: Transitive Agreement, Participant Objects (B forms)

S ↓ O →	1-	1+	12	12+	2-	2+
1-	-	-	-	-	wa	ŋanamba
1+	-	-	-	-		
2-	numba *or* nimba		-	-	-	-
2+f			-	-	-	-
2+m			-	-	-	-
2+pl	ŋiri		-	-	-	-

f	ŋa-m-ba-ŋi 1.INV.B.f	na-m-ba-ŋi 1+.INV.B.f	na-m-ba-ŋi 12.INV.B.f	ŋa :-m-ba-ŋi 12+.INV.B.f	a-[N]-ŋi 2-B.INV.f	na-m-ba-ŋi 2.INV.B.f
m	ŋa-m-ba-ni 1.INV.B.m	na-m-ba-ni 1+.INV.B.m	na-m-ba-ni 12.INV.B.m	ŋa :-m-ba-ni 12+.INV.B.m	a-[N]-ni 2-B.INV.m	na-m-ba-ni 2.INV.B.m
+	ŋa-m-bam-bi 1.INV.B.+	na-m-bam-bi 1+.INV.B.+	na-m-bam-bi 12.INV.B.+	ŋa :-m-bam-bi 12+.INV.B.+	a-m-bi 2-B.INV.+	na-m-bam-bi 2.INV.B.+
A	ŋa-m-baŋ-gu 1.INV.B.A	na-m-baŋ-gu 1+.INV.B.A	na-m-baŋ-gu 12.INV.B.A	ŋa :-m-baŋ-gu 12+.INV.B.A	a-ŋ-gu 2-B.INV.A	na-m-baŋ-gu 2.INV.B.A
M	ŋa-m-ba-ma 1.INV.B.M	na-m-ba-ma 1+.INV.B.M	na-m-ba-ma 12.INV.B.M	ŋa :-m-ba-ma 12+.INV.B.M	a-[N]-ma 2-B.INV.M	nam-ba-ma 2.INV.B.M

TABLE IV: Transitive Agreement, Nonparticipant Objects (B forms)

S ↓ O →	f	m	+pl	A
1	ŋa-[a]-ŋu 1.B.f	ŋa-[a]-nu 1.B.m	ŋa-m-ba 1.B.+	ŋa-[a]ŋ-gu 1.B.A
1+f		na[:]-a-ŋi-nu 1+.B.f.m	na[:]-am-ba-ŋi 1+.B.+.f	na[:]-a-ŋi-ɰu 1+.B.f.A
1+m			na[:]-am-ba-ni 1+.B.+.m	na[:]-a-ni-ɰu 1+.B.m.A
1+pl	na[:]-a-ŋu 1+.B.f	na[:]-a-nu 1+.B.m	na[:] -am-ba 1+.B.+	na[:]-aŋ-gu 1+.B.A
12	na-[a]-ŋu 12.B.f	na-[a]-nu 12.B.m	na-[a]m-ba 12.B.+	na-[a]ŋ-gu 12.B.A
12+f		ŋa[:]-a-ŋu-nu 12+.B.f.m	ŋa[:]-am-ba-ŋi 12+.B.+.f	ŋa[:]-a-ŋi-ɰu 12+.B.+.A
12+m			ŋa[:]-am-ba-ni 12+.B.+.m	ŋa[·]-a-ni-ɰu 12+.B.m.A
12+pl	ŋa[:]-a-ŋu 12+.B.f	ŋa[:]-a-nu 12+.B.m	ŋa[:]-am-ba 12+.B.+	ŋa[:]-aŋ-gu 12+.B.A
2	ba-ŋu 2-B.f	ba-nu 2-B.m	ba-m-ba 2-.B.+	ba-ɰu 2-B.A
2+f		nu -m-bu-ŋu-nu 2.B.+.f.m	na-m-ba-V -ŋi 2.B.+.+.f	ni-m-bi-ŋi-ɰu 2.B.+.f.A
2+m			na-m-ba-V -ni 2.B.+.+.m	ni-m-bi-ni-ɰu 2.B.+.m.A
2+pl	ni-m-bi-ri-ŋa 2.B.+.pl.f	nu-m-bu-nu 2.B.+.m	na-m-ba-[V]-ra 2.B.+.+.pl	ni-m-bi-ri-Ø 2.B.+.pl.A
f	wa-ŋu B.f	wa-ŋu-nu B.f.m	wam-ba-ŋi B.+.f	wa-ŋi-ɰu B.f.A
m		wa-nu B.m	wam-ba-ni B.+.m	wa-ni-ɰu B.m.A
+f		wam-bu-ŋu-nu B.+.f.m	wam-ba-V-ŋi B.+.+.f	wam-bi-ŋi-ɰu B.+.f.A
+m			wam-ba-V -ni B.+.+.m	wam-bi-ni-ɰu B.+.m.A
+pl	wam-bi-ri-ŋa B.+.pl.f	wam-bu-nu B.+.m	wam-ba-[V]-ra B.+.+.pl	wam-bi-ri-Ø B.+.pl.A
A	wa-ŋi-ŋ-gu B.f.INV.A	wa-ni-ŋ-gu B.m.INV.A	wam-ba-ŋ-gu B.+.INV.A	Ø
M	wa-ŋi-[N]-ma B.f.INV.M	wa-ni-[N]-ma B.m.INV.M	wam-ba-[N]-ma B.+.INV.M	Ø

NOTES to TABLES I-IV

In general: Affixes are separated by hyphens [-]; below each parsable affix sequence appears a gloss. Affix glosses are separated by periods [.].

Notes to TABLE I: All 3rd person subjects acting on 1st or 2nd person objects contain an underlying Inverse marker -n- which is overtly visible only in the + and A subject forms. Elsewhere it assimilates to a following nasal and deletes, since Nunggubuyu permits no surface geminates. This disappearing INV is notated as [N]. See section 3.2.3 for discussion.

Notes to TABLE II: In all the 2+m and 2+f subject forms and when augmented subjects act on plural object, vowel length (notated -V-) after the first affix derives from underlying -wV-, the affix realizing the [+aug] feature of the subject. Vowel length is thus a separate affix in these forms. Because of a late morphophonological rule shortening all vowels before -rV- 'plural' this -V- does not appear in any plural subject forms on the surface, nor do normally long person markers such as nV: '1+' or ŋV: '12+' appear long with the plural affix. Furthermore, overlong /V:-V/ always shortens to /V/, so in this context underlying -wV- also does not appear. These disappearing long vowels are shown above in brackets as [:] and [V] Compare nu :-nu '1+m/pl : m' with vowel length vs. na[:]-ra '1+pl : +pl' with disappearing vowel length, and ŋu :-nu '12+ : m' with vowel length and ŋa[:]-V-ni '12+m : +' with overlong shortened to long. In Table I, when optional (ra) 'pl' does appear with 12+pl object, the preceding ŋa :- '12+' is regularly short.

Notes to TABLES III and IV: The B tense/mood marker has the underlying form /wan /. The initial /w/ disappears intervocalically but hardens to /b/ after a nasal. The final /n/ assimilates in place to a following consonant and deletes before another nasal. The vowel /a/ tends to also disappear in certain forms by (more or less) regular phonological processes, but sometimes exercises a regressive /a/-harmony in preceding affixes. Owing to these alternations, the B affix appears in the above tables in the following forms, where segments in parentheses do not surface: wam, bam, baŋ, wa, ba, am, a, [a]m, [a]ŋ, m, [a].

In Tables II and IV, all nonsingular 3rd person objects (+m, +f, +pl) neutralize to one [+pl] form, which is represented in the [+pl] column. NA [m*] and MA objects are omitted for reasons of space: these present no interesting complications.

object and subject categories which are combined into single rows or columns: for example, +f, +m, +pl all neutralize to one row-- [+] in Table I (participant object). Other neutralizations entail only that certain 'expected' affixes do not appear. For example, not all entries in the column for +pl object in Table II have the plural object morph -ra. The reader will find it helpful to refer to Tables I and II and the list of affixes in (32) while going through the following discussion.

I will treat these neutralizations in several parts. The first group of neutralizations represent ceilings on the complexity of the object argument. The second group comprises ceilings on the complexity of subject arguments when combined with particular object arguments. The last group consists of neutralizations of gender-class when both arguments are marked for these, and the enhancement of [+aug] by 'pl' in certain circumstances. All the neutralizations will be expressed by filters combined with the automatic deletion of the lower of offending features, according to the Impoverishment hypothesis presented in the Introduction.

3.2.2.1 Object Impoverishment

All 19 agreement categories (22 categories minus 3 owing to the [±human] neutralization cited above for the 3rd person genders) are represented as subjects, but only 11 for objects.[5] By part of a process which Heath (1984) calls 'Dual-Trial Absorption' the augmented object categories undergo significant reduction as follows:

(38) a. 1+m, 1+f, 1+pl --> 1+
b. 12+m, 12+f, 12+pl --> 12 + (optional for 12+pl)
c. 2+m, 2+f, 2+pl --> 2+

The affected categories include functional duals and trials as well as plurals (Heath neutralizes the plural with another rule; the present analysis differs in this respect). The generalization is that the participant categories cannot have gender when objects. This represents one ceiling on the markedness of the object category: person features cannot cooccur with gender features in the object.

Because this neutralization occurs irrespective of the subject category with which a given object category is paired, it can be expressed as an Impoverishment of the OBJ AGR alone:

---

[5]Actually there are 12, since [-human] NA (m*) differs from its [+human] (m) counterpart by vowel harmony, as can be seen in Table II.

(39) Participant-Object Gender Deletion (first part)

*[$_{\text{OBJ AGR}}$ +participant +aug GENDER] (optional if also [+I +you])
(hence GENDER --> Ø)

The second part of Dual-Trial Absorption is not a neutralization, but rather has the effect of a structure-changing syncretism. For object categories, 3 m dual (+m) and 3 f dual (+f) become plural gender. In other words, 3rd person functional duals (the nonplural augmented categories) become explicit plurals, in a manner analogous to the replacement of explicit dual by plural in Warlpiri, as discussed in section 3.0.1. For now, assume that the rule *converts* the gender of an augmented object to plural:[6]

(40) Dual-Trial Object Absorption

m, f --> pl / [$_{\text{OBJ AGR}}$ +aug ____]

As a result of rule (40), 3rd person dual objects behave morphologically (and morphosyntactically) as if they were plurals. (I will later show that this rule actually consists of two parts, a deletion of 'm/f' and an insertion of 'pl'.) Plural gender introduced by rule (40) figures in later rules of Impoverishment as I will show later. Moreover, although generally obscured, this plural gender surfaces in a few forms, in particular, by showing the plural gender affix -rV. Hence, for example, the following clitics serve both for plural and dual objects (a colon separates subjects from objects in the clitic cluster gloss, a dot between glosses of each constituent affix):

(41)

a.	ŋa-ra 1.pl	'1 sg: 3 dual / pl'
b.	na-ra 12.pl	'1 incl dual: 3 dual / pl'
c.	ba-ra[7] 2.pl	'2 sg: 3 dual / pl'
d.	wa-[V]-ra[8] +.+.pl	'3 pl : 3 dual / pl'

---

[6]Inasmuch as this rule requires reference to m and f to the exclusion of pl, the former two may be designated by a feature [-plural], whereas pl = [+plural]. This has little consequence for the general system, so for now I simply assume that gender is expressed by privative features.

[7]Although not in the B-form, the 2sg affix appears as /ba/ here by a special allomorphy rule.

[8]/w/ disappears intervocalically; a late morphophonological rule shortens any long vowel before -rV 'plural,' such that the augmented affix -wV- disappears entirely in this and other forms.

In each case, the plural affix -ra is the realization of the plural feature introduced by rule (40), and shows the [a]-harmony of an object morph. Even when collocated with an explicitly dual argument, the clitic is plural in form, for example:

(42) nu:- ˈbi-n̪i, na-nuŋ-gaɽagu-wa: , **wa-ra**-ŋuɽa=guldha-ŋi
m-ANAPH-m(dual) m-uncircumcised-dual aug-pl-penis=cut-$past_2$

those two, two young (uncircumcised) boys, they circumcised them (§2.1.2)

Finally, the last neutralization that affects object categories only is the deletion of [+aug] in 2+ object category:

(43) 2-Object Number Deletion
*[OBJ AGR 2 +aug] (cf.Heath's 'Nonsg-Deletion' p. 370)

As a result of (43), 2nd person augmented objects appear without the [+aug] affix -WV. An 'inverse' marker -n-, to be discussed later, also appears in these forms, whenever an object affix precedes a subject affix in left-to-right linear order:

(44) a. na-n-ni --> na-ni
2.INV.m '3 m sg: 2 dual / plur'

b. na-n-ɰu --> na-ŋ-gu
2.INV.A 'ANA : 2 dual / plur'

Second person singular objects remain distinct in virtue of having a specific [2-] affix which appears here as nu-, harmonizing as ni-:

(45) a. nV-n-ni --> ni-ni
2-.INV.m '3 m sg: 2 sg

b. nV-n-ɰu --> nu-ŋ-gu
2-.INV.A 'ANA : 2 sg'

3.2.2.2 Subject Impoverishment

A neutralization similar to (39) occurs in subject categories as well. A gender feature is deleted in augmented categories when the object is [+participant]:[9]

(46) Participant-Object Gender Deletion (second part)

*[$_{\text{SUB AGR}}$ +aug GENDER] [$_{\text{OBJ AGR}}$ +participant].

(cf. Heath's Plural Deletion d)

In virtue of (46), the 3rd person dual and plural subject arguments (+m, +f and +pl) neutralize to merely augmented (i.e. +) when the object is 1st or 2nd person. This entails that no gender affix will appear in the following combinations:

(47)

a. ŋa-n-wi --> ŋa-m-bi
1.INV.+ '3 m du/ f du / pl: 1 sg'

b. na-n-wi --> na-m-bi
1+.INV.+ '3 m du / f du / pl: 1 sg'

c. na-n-wi --> na-m-bi
12.INV.+ '3 m du / f du / pl: 12'

d. ŋa-(ra)-n-wi --> ŋa-(ra)-m-bi
12+.pl.INV.+ '3 m du / f du / pl: 1 incl trial / pl'

e. ni-n-wi --> ni-m-bi
2.INV.+ '3 m du / f du / pl: 2 sg'

f. na-n-wi --> na-m-bi
2+.INV.+ '3 m du / f du / pl: 2 dual / pl'

A textual example of the last form (47f) is shown below. (This also occurs with the B tense/mood affix -bam after INV, but the point is that neither 'f' nor 'pl' appears in the sequence):

(48) na-m-bam-bi = ṛa-yi :
2+.INV.B.+ = spear

'They will spear you (two women).' (§17.12.3)

---

[9]The only exception to this rule may be that when the subject is 2+pl, it seems to retain the 'pl' -rV- when the object is 1st person in the clitic ŋi-ri '2 dual / pl: 1 sg/ excl pl.' However, this form is probably learned as an unanalyzable portmanteau conditioned by [$_{\text{SUBJ}}$ 2 aug][$_{\text{OBJ}}$ 1]. See also fn. 7.

In addition, (46) causes 1+m, 1+f, and 1+pl subjects to neutralize to merely 1+ when the object is [+participant], *de facto* 2. A further Impoverishment also applies in this circumstance:

(49) 1-2 Number Deletion

*[+I +aug] [+you]

(cf. Heath's 1-->2 and 2 --> 1 Neutralization, p. 364)

When both arguments are [+participant], i.e '1:2' or '2:1,' the 1st person argument cannot also be [+aug], either when it is subject or object. The [+aug] feature is deleted by (49), and the category merges with the 1 sg argument. Observe that in the following combinations, there is no distinction among 1 sg and 1 dual or 1 pl:[10]

(50)

a. ŋu-nu
1.2- '1sg/1excl dual or plural : 2 sg' = 'I:thou *or* we:thou'

b. ŋa-na
1.2 ''1sg/1excl dual or plural : 2 dual / pl' = 'I:you *or* we:you'

c. ŋi-ri
1.pl '2 dual / pl: 1 sg/ excl pl'[11] = 'you: me *or* you: we'

What is significant about the cooccurrence restriction (46), which entails the deletion of the GENDER feature, is its similarity to the restriction in (39) for object categories. The facts are roughly the same: for objects, GENDER is not permitted for [+participant] categories, which are necessarily [+aug] in virtue of cooccurrence restriction (35); likewise, when the object is [+participant], the subject cannot have gender either, when it is also [+aug]. To unite the two cases, we must envision the following process: when SUBJ and OBJ agreement heads merge into a single M^0, the resulting fused 'morpheme' has the same internal structure as each of its constituent parts. Thus, a cooccurrence restriction holding of one of the parts may also hold of the fused M^0 itself. To see this, consider the two forms below and the Impoverishments they undergo, with a relevant clitic shown in (52). In both cases the gender feature 'm' is deleted:

[10]The B form analogues to (50a,b), namely wa- and ŋanamaba- are not immediately parsable and probably learned as unanalyzed portmanteaus.

[11]The '2:1' combinations are subject to further irregularities which will not be discussed here: it is likely that they are learned as indivisible portmanteau forms. This form is provided here only because it obeys the rule, although it is probably a portmanteau in the synchronic grammar of Nunggubuyu.

(51) a. [OBJ AGR 1+m] --> [OBJ AGR 1+] (by (39))

b. [SUB AGR +m][OBJ AGR 1] --> [SUB AGR +][OBJ AGR 1] (by (46))

(52) a. na-n- ŋi --> na-ŋi
1+.f '3 f sg: 1 excl m dual' (i.e. 'f: 1+m')

b. ŋa-n-wi --> ŋa-m-bi
1-.INV. + '3 m du : 1 sg'

In (52a) the masculine gender of the object is Impoverished, so that in the resultant sequence of affixes, the object is represented only by the prefix na- '1+.' Similarly in (52b), the masculine gender of the subject is Impoverished when the object is [+participant], here 1 sg. The resultant sequence has no gender affix.

From this it is possible to conclude that merged M^{0}s share certain of the same complexity thresholds as do single AGR affixes. This allows a straightforward explanation for the next set of neutralizations.

3.2.2.3 Gender Impoverishment

Whenever the subject and object arguments bear the *same* gender feature, the subject feature is deleted:

(53) *[SUB AGR αGENDER] [OBJ AGR αGENDER]

(cf. Heath's Plural Deletion a, Gender Marker Deletion a, b)

This constraint mirrors the same constraint operative on single categories: a single category cannot be doubly masculine, doubly feminine, or doubly plural. The effect is as if when merging two AGR heads, where two identical features come to reside in the same M^0, they are capable of being realized as only one affix. Recall that we saw the same phenomenon occur with 'pl' in Maung in section 3.0.3.

As a result of (53), where both arguments are masculine, feminine, or plural, only one gender marker-- that of the object -- will appear, e.g.:

(54) a. nu (not *nu-nu)
m '3 m sg: 3 m sg'

Ex: nu- maraŋ = a-ñ 'he handed (it) to him' (§4.6.4)

b. nu :-nu (not *nu :-nu-nu)
1+.m '1 excl dual m: m'

c. ŋu (not *ŋu-ŋu)
f '3 f sg: 3 f sg'

Ex: ŋu-waraga = wagiwa-ñ
f - upper back = hit - past$_1$ 'she hit her in the upper back' (§6.3.3)

d. wa-[V]-ra (not *wi-ri-ra)
+.+.pl '3 pl : 3 pl'

e. na-[V]-ra (not *ni-ri-ra)
2.+.pl '2 pl: 3 pl'

The reason that it is the *object* feature and not the subject feature which wins out follows either from the order of embedding of objects and subjects or from general principles of markedness. Objects will be lower syntactically than subjects and therefore will be more embedded in the cluster. Under the markedness approach, in a nominative-accusative system, objects are more marked morphosyntactically than subjects (cf. accusative and nominative case marking). The evidence that the object is the winner comes from allomorphy and ablaut. Objects 'pl' affixes show an [a]-harmony allomorph -ra, and this allomorph appears in forms neutralized by (53). Examples are above in (54d,e). Also, the [-human] masculine (NA class) shows [i]-harmony when the object, as opposed to the normal [u]-harmony . When a masculine subject acts on [-human] masculine and (53) applies, the [i]-harmonic morph appears:

(55) ni-
m* '3 m sg human > 3 m sg nonhuman'

Ex: lhawulgulg ni = wi-ni
sharp-bladed-grass m* = hit

'He hit it (sharp bladed grass, m*) (§18.19.2)

This shows that it is indeed the object feature which survives the filter in (53).

Beyond this relatively simple neutralization, there exist several other neutralizations governing the cooccurrence of gender in transitive forms. These can be schematized below, with examples given in (57):

(56) a. m --> Ø / [$_{\text{OBJ AGR}}$ f] (cf. Heath's Gender-Marker Deletion a)

b. pl --> Ø / [$_{\text{OBJ AGR}}$ m] (cf. Heath's Plural Deletion b)

c. pl --> Ø / [$_{\text{OBJ AGR}}$ ____] [$_{\text{SUBJ AGR}}$ GENDER / CLASS]

(57) a. ŋu (not *ni-ŋu)
f '3 m sg: 3 f sg'

Ex.: ŋu - lhaŋarma : -
f - reach - $past_2$ 'he caught up to it (python, f*)' (§1.6.2)

b. wu-nu (not *wu-ru-nu)
+.m '3 pl: 3 m sg'

Ex: wu-nu = ya-ñ
+. m = give - $past_1$ 'they gave (him) to him' (§5.17.6)

c. na[:]-a-ni
1+.+.m '1 m trial: 3 pl'

Rules (56a,b) continue the impoverishment of subject gender by object gender; rule (56c) however, deletes object 'pl' whenever the subject has gender or class features. This rule is bled by (53) in the case where both subject and object are plural: here the object wins, as in (54d,e).

If feature hierarchies determine the direction of neutralization, then rules (56a,b) suggest the hierarchies f > m and m > pl, since a feminine object deletes a masculine subject by (56a) as in (57a), and a masculine object deletes a plural subject by (56b) as in (57b). Thus, these rules can be understood as subcases of a more general process of (53), where rather than strict identity of feature, what is required is that the object gender be higher than or equal to the subject gender in the hierarchies f >m and m > pl. We may then rewrite (53) as (58):

(58) Gender Impoverishment

*[$_{\text{SUBJ AGR}}$ αGENDER] [$_{\text{OBJ AGR}}$ βGENDER]
where hierarchically $\beta \geq \alpha$

By transitivity of the hierarchy, we should also expect f > pl. But there is in fact a curious failure of transitivity: it is not the case that a feminine object deletes a plural subject, since these may cooccur, as in:

(59) a. ni [:]-ri-ŋa
1+.pl.f ‘1 excl dual / plur: 3 f sg’

Ex: ni [:]-ri-ŋa-wulgu = wulguldha-ŋi
1+.pl.f = cut up ‘we (Nunggubuyu people) cut it (palm, f*) up’
(§117.6.1)

b. ŋi [:]-ri-ŋa
12+.pl.f ‘1 incl pl: 3 f sg’

These are the actually occurring forms. If the ‘f’ object feature were ‘strong enough’ to delete the subject plural feature, then the forms would be the following (cf. the well-formed masculine object forms with suffix -nu ‘m’ instead of -ŋu ‘f’):

(60) a. *nu :-ŋu
1+.f

b. *ŋu :-ŋu
12+.f

It appears, therefore, that Gender Impoverishment (58) is constrained to apply only if the relevant features are *adjacent* on the hierarchy. The hierarchy, therefore, should be understood as the graph below, where, if feature β points at feature α, then β when part of OBJ AGR is ‘strong’ enough to delete α when in SUB AGR:

(61) Gender Hierarchy Graph

$$\begin{tikzcd} f \arrow[r] \arrow[loop above] & m \arrow[r] \arrow[loop above] & pl \arrow[loop above] \end{tikzcd}$$

All features points to themselves, so all object features delete identical subject gender features (cf. filter (53). Additionally, ‘f’ deletes ‘m’ and ‘m’ deletes ‘pl,’ as stated in rules (56a,b).

What is special about this graph is that it is not *connex*, i.e. not every point enters into a direct relation with every other point. In particular, ‘f’ and ‘pl’ do not interact. Whether feature hierarchies in general permit such non-connex relations or whether this is an accidental fact about Nunggubuyu remains an open question.

Turning finally to rule (56c), a 'pl' object feature deletes whenever the subject has a gender or class feature. This rule does not obey any hierarchy, since we should expect object gender features to 'beat out' subject gender features. As such, it is desirable to express this process in terms other than Impoverishment, which is constrained to obey the hierarchy of features. I now explain how this is accomplished.

The feature 'pl,' which is the target of rule (56c), is espe ..ly recessive within the morphology of Nunggubuyu, because it serves as an enhancement of [+aug], as was mentioned in section 3.0.1. Any augmented category not marked for masculine or feminine gender can be predictably assigned 'pl.' (If the category is already 'm' or 'f' then this rule cannot apply, since each category may have only one gender feature):

(62) Plural Enhancement

[+aug] --> pl

This enhancement is constrained to apply only when it does not violate cooccurrence restrictions such as (39) and (47). Given this enhancement rule, it is also now possible to replace rule (56c) by a cooccurrence restriction as below:

(63) Plural Object Restriction

*[$_{\text{SUB AGR}}$ GENDER/CLASS] [$_{\text{OBJ AGR}}$ pl]

In virtue of this restriction, (62) will be unable to insert 'pl' into an object category if the subject is marked for gender or for class. This will have the same effect as rule (56c).

When both subject and object are 'pl,' a complication arises. Recall that Gender Impoverishment bleeds (56c). On the enhancement analysis, we must assume that (62) applies to the object first, supplying 'pl' to it, before it supplies 'pl' to the subject. If the reverse order were to hold, then a 'pl' supplied to the subject would suffice to block the insertion of 'pl' to the object because of (63). The 'pl' object feature does appear, however (review (54 d,e). In this case, it is a natural assumption that default enhancement rules such as (62) apply to the inner (object) constituent first before applying to the outer (subject) constituent, giving the correct ordering of events.

If 'pl' is automatically inserted as an enhancement of [+aug] by rule (62), it no longer becomes necessary to assume a structure-changing syncretism for Dual-Trial Object Absorption, rule (40), either. The old version is repeated below:

(64) Dual-Trial Object Absorption

m, f --> pl / [OBJ AGR +aug ____]

Instead, the process above can be reanalyzed as the deletion of 'm' or 'f,' (65a) followed by the enhancement of an [+aug] category by 'pl' (65b):

(65) Dual-Trial Object Absorption (Revised)

a. *[OBJ AGR +aug m/f] (hence m/f --> Ø)

Plural Enhancement

b. [+aug] --> pl = (62)

This allows the rather stipulative structure-changing rule (64) to be replaced by an Impoverishment (65a) obeying the hierarchy of features, and the independently-needed Plural Enhancement.

The number of circumstances in which Plural Enhancement can apply is small. If inserting 'pl' into the object, the subject must not have any gender or class feature. This limits its application to singulars and to plural subjects (whose 'pl' feature has not yet been inserted, and never will be), for example:

(66)

a. ŋa-[V]-ra
1.pl — '1 sg: 3 pl'

b. na[:]-[V]-ra
1+.+.pl — '1 excl pl: pl'

If inserting 'pl' into a subject, Plural Enhancement can apply only in three cases. If the object is [+participant], 'pl' cannot be inserted into the subject owing to (46). When the object is [-participant], it can apply only when not in violation of Gender Impoverishment (58), i.e. only when the object is 'f' or CLASS, for example:

(67)

a. wi-ri-ŋa
+.pl.f — '3 pl : 3 f sg'

b. wi-ri-Ø
+.pl.A — '3 pl: ANA'

c. wi-ri-ma
+.pl.M — '3 pl: MA'

To summarize the discussion of gender neutralizations, I have shown that when two gender features come together in transitive agreement, certain neutralizations take place. First, if both features are identical, then the object feature 'wins' and the subject feature is deleted. Furthermore, if the features are not alike, then the object gender feature deletes the subject gender feature if it is higher on the hierarchy f > m > pl, where the hierarchy is not a connex relation.

I argued that the particularly recessive nature of the value 'pl' within the transitive agreement system arises from two interacting factors. First, 'pl' is not underlyingly represented but is rather an enhancement of [+aug], which is inserted by a redundancy rule when an augmented category has no other marked gender feature. Second, a number of cooccurrence restrictions constrain this enhancement to apply only in a limited range of circumstances, where insertion of 'pl' will not surpass complexity thresholds.[12]

3.2.2.4 Summary of Neutralizations in A forms

The following presents a summary of the proposed Impoverishments, Syncretism and Deletion Rules, and Hierarchy of Gender Features. These seven constraints and rules derive all the neutralizations seen in the A forms:

(68) Rule/Constraint Summary

1. Dual-Trial Object Absorption (Revised)
   *[$_{\text{OBJ AGR}}$ +aug m / f] (hence m/f --> Ø)

2. Plural Enhancement
   [+aug] --> pl

3. Participant-Object Gender Deletion
   a. *[$_{\text{OBJ AGR}}$ +participant +aug GENDER] (optional if [+I +you])
   b. *[$_{\text{SUB AGR}}$ +aug GENDER] [$_{\text{OBJ AGR}}$ +participant]
   (hence GENDER --> Ø)

---

[12]There are as yet one more minor neutralization of gender-class. When both arguments are nonhuman, then the prefix is Ø-. Heath suggests that this may be the 'mutual destruction' of two affixes. Rather than see this as 'mutual destruction,' I prefer Heath's alternative proposal whereby when both subject and object are nonhuman, there is a special contentful Ø- realization for this combination.

4. 2-Object Number Deletion

$*[_{\text{OBJ AGR}}\ \text{+you +aug}]$ (hence +aug --> Ø)

5 1-2 Number Deletion

$*[\text{+I +aug}]\ [\text{+you}]$ (hence +aug --> Ø)

6. Gender Impoverishment

$*[_{\text{SUBJ AGR}}\ \alpha\text{GENDER}]\ [_{\text{OBJ AGR}}\ \beta\text{GENDER}]$ where hierarchically $\beta \geq \alpha$

(hence αGENDER --> Ø)

7. Plural Object Restriction (constrains Plural Enhancement)

$*[_{\text{OBJ AGR}}\ \text{pl}\ ]\ [_{\text{SUBJ AGR}}\ \text{GENDER / CLASS}]$

(69) Gender Feature Hierarchy Graph

f ---> m ---> pl

These affixes which actually appear in a form are those which realize features that survive Impoverishment . In the next section I detail how these affixes are Placed in Nunggubuyu. We will see that many of the above Impoverishments are motivated by the need to suppress the number of affixes that are realized so that generated strings can be correctly ordered by principles of hierarchy and by the merger relation identified in Marantz (1988).

3.2.3 Placing in Nunggubuyu

Once the Impoverishments discussed in the last section have applied to input morphosyntactic representations, the Placing of affixes within the clitic cluster proceeds by means of two ordering relations, one following from the hierarchy of features, the other from syntactic nestedness. I will show that in circumstances where the feature-hierarchy ordering is at variance with the syntactic ordering, a repair operation occurs, which has the effect of affix metathesis. Using the theory of Marantz (1988), I reduce this operation to that of second-position clitic merger, thereby explaining why only one affix may be moved by this operation, which is strictly local. I then show that, of the seven impoverishments discussed in the previous section, four can be motivated automatically as strategies to ensure that a generated string is repairable by merger.

The basic ordering relation can be summarized as a transitive relation among the affix types according to the features they discharge as below:[13]

(70) [+participant] > [+aug] > GENDER > CLASS

For example, the sequence for 2+pl acting on MANA is:

(71) ni-V-ri-ma
2.+.pl.M ‘2 pl: MANA’

When two arguments possess features of the same type in (70), then the subject affix will precede the object affix (cf. Heath's ‘Main Ordering Rule’, 1984: 367):

(72) Subject > Object

Examples include:

(73) a. ŋu-nu
f.m ‘3 f sg: 3 m sg’

b. ni-[V]-ri-ŋa
2.+.pl.f ‘2 pl: 3 f sg’

c. wa-V-ŋi
+.+.f ‘3 f dual : 3 nonsg’

d. ŋu-nu ‘1 sg/pl : 2 sg’
1 . 2-

The two ordering principles (70) and (72) often are at odds. This occurs whenever an affix indexing an object feature *precedes* an affix indexing a subject feature by (70). In

[13]The only exception to this rule is the form ŋa-ra-m-bi ‘3 nonsg : 12+pl.’ Here the affix -bi, marking an augmented subject, appears after the affix -ra marking a plural object. However, it is precisely in this case -- where the object is 12+pl, that the affix -ra is optional. This ordering, I surmise, reflects an older grammar where a more complicated set of principles was at work in ordering clitics (probably more consistent with Heath's analysis).

Another explanation is that the sequence ŋa-ra has been reanalyzed as a single portmanteau with the hierarchical status of a [+participant] affix. Forms with ŋa-ra object are also exceptional because they are the only ones in which two object affixes precede inverse -n-. According to the analysis to be presented here, this situation can only arise if ŋa-ra is an atomic constituent.

other words, it is not the case that *all* subject affixes must precede all object affixes; subject affixes precede object affixes only when they are of the same type. The ordering relation in (70) 'wins,' but at the juncture between the object affix and the subject affix, a special inverse marker -n- is inserted:

(74) **Inverse Insertion**

Ø --> [-n-$_{\mathrm{INV}}$] / $_{\mathrm{OBJ}}$] ____ [$_{\mathrm{SUB}}$

Usually this -n- deletes before a following consonant, but it does surface when the following segment is /w, ɰ / which harden to / b g / :

(75)	a. ŋi-ɰu --> f_{OBJ}.A_{SUB}	ŋi-n-ɰu --> ŋi-ŋ-gu f..INV.A	'ANA : 3 f sg'
	b. wa-ɰu --> $+_{\mathrm{OBJ}}$.A_{SUB}	wa-n-ɰu --> wa-ŋ-gu +.INV.A	'ANA : 3 nonsg'
	c. ŋa-ɰu --> 1_{OBJ}.A_{SUB}	ŋa-n-ɰu --> ŋa-ŋ-gu 1.INV.A	'ANA : 1 sg'
	b. ŋa-wi 1_{OBJ}. $+_{\mathrm{SUB}}$	ŋa-n-wi --> ŋa-m-bi 1.INV.+	3 nonsg: 1 sg'

In these examples, an object affix precedes a subject affix in virtue of (70). The resultant order is at variance with (72), and so inverse -n- appears manifestly at the 'offending' boundary. In other cases, inverse -n- is deleted by regular phonological rules, but Heath proposes (p. 85, p.362 'V-ablaut') that it also conditions a morphophonological [a]-harmony in preceding affixes, e.g.:

(76)	a. na-a-[V]-[n]-ŋi 2.+.+.INV.f	'2 f dual : 3 nonsg'
	b. ni- [V]-ri-ŋa 2. +. pl. f	'2 nonsg: 3 f sg'

In (76a), the inverse marker is associated with [a]-harmony in the preceding affixes; in (76b), without inverse, there is no [a]-harmony.

We now must explain why there should be two ordering principles, and how precisely the ordering is effected, in particular the unusual inverse forms.

First consider the ordering of Subjects > Objects. This ordering derives from the syntactic nestedness of the input cluster. Although Nunggubuyu is arguably one of the more radical examples of 'nonconfigurational' structure (Heath 1986), what evidence does exist for a hierarchical syntax derives from the highly articulated internal structure of the verb word. In particular, I will assume that the agreement clitic cluster represents the incorporation of the actual argumental projections in syntax, and that overt nominals are adjoined to the clause (Jelinek 1984). On this assumption, the object or patient argument must be syntactically lower than the agent or subject argument, by standard assumptions, to derive proper assignment of thematic relations. Therefore, when the verb incorporates into these arguments, the object will incorporate first ([*] represents left-to-right adjacency):

(77) $[_{D}$ SUB * $[_{D}$ OBJ [V ...]

This derives the basic Subject > Object ordering relation.

Now consider the second ordering principle (70). This principle is, in itself, quite simple. The affixes appear in the left-to-right order as motivated by the hierarchy of features. Person features [+participant] are higher than number features [+aug], which are in turn higher than gender and then class features. This is the same hierarchy which governs Impoverishment and morphological rule ordering. It is significant that the ordering is not, for example, gender > person > class > number. Such an ordering would not follow by principles of hierarchy and would, if it indeed ever occurs, be highly costly, owing to the explicitly stipulated ordering statements required to generate it.

On each cycle then we can expect the ordering [p > n > g > c] where 'p' is person, 'n' is number, 'g' is gender, and 'c' is class, giving, schematically:

(78) [p > n > g > c] * [p > g > n > c].

The result of the hierarchy is that *within each cycle* the affixes come to be ordered according to the principle in (70). However, we have already seen that in the inverse forms, e.g. (75), the ordering of (70) violates the Subject > Object principle. This entails that the ordering of affixes requires that the principle in (70) holds over the output of both cycles. As things now stand, the inverse forms in (75) will be generated as the forms in (79):

(79) a. [ɰɥɯ * [$_{OBJ}$ ŋi]]
A. f

b. [ɰɥɯ * [$_{OBJ}$ wa]]
A.+

c. [ɰɥɯ * [$_{OBJ}$ ŋa]]
A . 1

d. [wi * [$_{OBJ}$ ŋa]]
+ . 1

However, observe that the strict adjacency [*] relation in (78) holds over the whole subject and object constituents, and not of the *individual* affixes contained within them. The ordering principle (70) requires only that each individual p precede each individual n, which must precede each individual g, and so forth. I now propose that the inverse forms arise when a single object affix merges with the final subject affix (in the sense of Marantz 1988). To see this, let us represent the last two affixes of the subject cluster as A and B and the first affix of the object cluster as C:

(80) [[..... A] * B $_{SUBJ}$] * [$_{OBJ}$ C]

Now suppose that C must precede B by the ordering relation in (70), for example if B is a gender affix ni 'm' and C is a number affix -[W]V- 'augmented.' To repair this, C can 'merge' with B, replacing the adjacency requirement by the merger relation. In other words, it is possible to treat C is as a second-position clitic (here, second to last, within the clitic cluster). Because B is the last affix to be added within the subject clitic cluster, it is the head of that constituent, and therefore the host of C.

It is precisely in this circumstance that the inverse marker -n- appears. This marker, I propose, is introduced as a signal of the merger relation:

(81) [.... A] * [C -n- B]

Now an interesting claim results from this analysis of the inverse forms. Returning to (80), it is the case that the rebracketing and merger can take place only when the object contains only one affix. To see this consider the situation when the object has more than one affix:

(82) [[... A] * B] * [[C] * D]

In (82), C cannot merge with B since C must remain adjacent to D: C has two adjacency requirement and so metathesis via merger cannot take place (cf. Marantz 1988: 265). It follows, then, that in all cases, only one affix will follow the inverse marker -n- as in (81). This prediction is correct: all the clitic clusters in Nunggubuyu have but one affix following -n-. It is important to observe that if inverse -n- were inserted only whenever an object affix preceded a subject affix, there would be no reason to expect this generalization. On the present analysis, where inverse always marks merger, the fact that inverse is followed by at most one affix follows by absolute necessity. (The 'B' affix to be discussed in the next section may also follow inverse, but is inserted after this merger).

If an object has more than one affix, of which at least one violates the ordering principle (70), then merger is not a possible repair strategy to achieve well-formedness. Thus, only those combinations leading to violations of (70) in which the object has been sufficiently Impoverished (down to one affix) can result in well-formed clitic sequences. This consideration, it turns out, motivates four of the Impoverishments discussed in the last section.

First, consider the two impoverishments Participant-Object Gender Deletion (part a), and 2-Object Number deletion:

(83) * [$_{\text{OBJ AGR}}$ +participant +aug GENDER]
*[$_{\text{OBJ AGR}}$ +you +aug]

By deleting the gender of any [+participant] object, and the number of any [+you] object, these impoverishment rules guarantees that, whenever the form is inverse, the object sequence will have at most one affix. For example, consider the following input:

(84) [$_{\text{SUB AGR}}$ f] [$_{\text{OBJ AGR}}$ 1 + m]

If the object is not impoverished as to gender, the resultant string will be:

(85) ŋi * [[na] * nu]]

The affix na '1+' must precede ŋi 'f' by (70), but, because of its adjacency requirement to nu 'm', it cannot 'escape' its position by merging with ŋi to give *na-n-ŋi-nu. Thus, any form in which 'm' is not impoverished will be unable to generate a well-formed string.

Similarly, [+aug] must be impoverished in the 2-object forms, since in the 2nd person, [+aug] is realized as a separate affix. For example, consider the input

(86) $[_{\text{SUB AGR}}\ \text{f}\,]\,[_{\text{OBJ AGR}}\ 2+\,]$

Without impoverishment of [+aug] in the object, this will yield the string

(87) ŋi * [[na] * wV]]

Both na and wV must precede ŋi by (70), but na cannot escape to merge with ŋi as long as it must remain adjacent to wV. Therefore, the only derivation which will give a well-formed clitic sequence is one in which [+aug] is impoverished from a [+you] object. (The first person forms do not present a problem, since [+aug] is realized together with the person affix and is not a separate affix.)

Similarly, Plural Object Restriction also impoverishes the object sequence to one affix:

(88) $*[_{\text{OBJ AGR}}\ \text{pl}]\,[_{\text{SUB AGR}}\ \text{GENDER/CLASS}]$

Consider a situation if Plural Object Restriction has not applied, for example:

(89) $[_{\text{SUB AGR}}\ \text{f}\,]\ *[_{\text{OBJ}}\ +\text{pl}]$

↓

[ŋi] * [[wV] * [ra]]

Here, because wV 'augmented,' a number affix, must precede ŋi 'f,' a gender affix, then it must metathesize to the left of ŋi. But it cannot escape owing to the following affix ra. Therefore, a well-formed string can be generated only when 'pl' is impoverished from the representation, leaving only one object affix.

A subject feature impoverishment is the fourth one motivated by this same consideration. Because merger must metathesize a single object clitic with the *final* affix of the subject, this final affix must be the only one within the subject affix sequence which must follow the object affix by (70). The relevant situation is schematized below:

(90) $[[\text{n}]\ *\ \text{g}_{\ \text{SUB}}]\ *\ [_{\text{OBJ}}\ \text{P}$

Since p must precede both n and g, the string in (90) cannot be adequately repaired by merger, since p can merge only with g, whereupon it will still be to the right of n. Therefore, when the object is [+participant], the subject cannot have both a number and a gender affix. This is effected by Participant Object Gender Deletion (part b), repeated below:

(91) *[$_{\text{SUB AGR}}$ +aug GENDER] [$_{\text{OBJ AGR}}$ +participant]

To summarize, the Placing of affixes in the Nunggubuyu clitic cluster obeys two principles. One places all Subjects affixes to the left of Object affixes. I proposed that this placement derives from syntactic nestedness. The second principle orders affixes within these constituents according to the hierarchy of features, namely, person before number before gender before class. This second principle holds then of the entire sequence of both subject and object clitics. The merger operation (Marantz 1988) applies when possible to repair orders which offend the second principle (cf. Heath's 'Gender Hopping' rule, p. 368). The inverse marker -n- is inserted by rule wherever merger takes place.

Because merger can apply only to peripheral object affixes, I explained why inverse -n- is always followed by at most one affix. I then showed that Participant-Object Gender Deletion (two parts), 2-Object Number Deletion and Plural Object Restriction (Impoverishments discussed in the previous section) are all motivated to prevent the spelling-out of strings of affixes whose ordering could not be repaired by the merger operation. This provides an entirely natural and system-dependent explanation for why these and not other Impoverishments occur. Indeed, a further analysis of the system of neutralizations (not attempted here) may reveal a similar motivation for (at least part of) the Gender Impoverishment facts as well, but I leave this an open matter.

3.2.4 Neutralization and Placing in the B forms

Tables III and IV (pp. 253-54) present the transitive agreement forms which arise in certain tense/mood combination which Heath calls the B forms. The precise morphosyntactic distribution of the B forms is somewhat complex (Heath 1984:338-39). B forms occur in the negative past and present, the affirmative future, and always in the past potential. As such, the B affix is conditioned by tense-mood features, although tense-mood also has a separate suffixal position of exponence. Thus, every Nunggubuyu verb contains at least the following (cf. Heath 1984:145):

(92) [(B) + agreement clitics] = Verb - INFL

Between the B affix plus agreement clitic cluster there is a major phonological boundary, notated above by the sign =.

One typologically unusual aspect of Nunggubuyu is that it has no complementizers, words corresponding to English *that* or French *que*, and indeed, Nunggubuyu lacks any nonfinite complement clause types (Heath 1984:558). Assuming that the complementizer functional category is universally available, then the B affix may in fact be a type of complementizer, which attaches peripherally to the clitic sequence. If the B-affix is in fact a cliticized Comp, then this Comp can show concord or a selectional relation with Tense/Aspect and Negation (cf. Laka 1990); thus these morphosyntactic properties can be manifested twice in the complete verb word. Although this proposal is somewhat conjectural, I will take as a hypothesis the idea that the B affix originates from a syntactic source external to the Verb+INFL and attaches to the clitic sequence.

Observationally, combinations of B (or its absence) and INFL are correlated with the Tense-Aspect-Mood-Negativity categories. For example, allomorph 2 of the past INFL is correlated with the Past Actual Continuous, but when combined with the B form, it is correlated with the Negative Past Actual (Heath 1984:338).

Morphologically, the B forms present a number of minor complications which will be addressed in this section.

The B-form affix is underlyingly wan. The initial /w/ hardens to /b/ after nasals, disappearing intervocalically and optionally in initial position. The final /n/ assimilates to the place of any following consonant, and deletes before a following nasal by geminate reduction. These phonological processes are entirely regular within the clitic cluster, being evidenced elsewhere in the alternations of wa 'augmented' and nun '2.' After ŋa- '1' and na- '12', underlying wan sometimes disappears completely instead of lengthening the preceding prefix (see 1 and 12 rows in Table IV). This is consistent with a general rule which deletes all vowels after these prefixes (cf. rows 1 and 12 in Table II, where [V] deriving from wa '+' also deletes).

Thus, depending upon its phonological environment, the B affix has one of the following surface instantiations, where a segment in brackets deletes by regular rules: ba, baŋ, bam, [a], a, m, wa, wam, am, aŋ, [a]ŋ.

The B affix is Placed directly after the first affix realizing person features; if there is no person affix, then the B affix is initial:

(93) a. ni-*m*-bi-ŋi-ɰɟu < ni-wan-bi-ŋi-ɰɟu
2.B.+.f.A '2+f : A [B form]' 'you (f dual) : it (ANA gender)'

b. *wam*-bi-ŋi-ɰɟu
B.+.f.A '+f : A [B form]' 'they (f dual) : it (ANA gender)'

In (93a), the B affix appears after the person affix ni- '2.' In (93b) there is no affix realizing person, since both arguments are 3rd person. The B affix is therefore initial.

The 2nd person unaugmented (2-) affix has a special allomorph for the B form, namely ba- (or a- before Inverse). With this special allomorph, the independent B affix does not appear (see the 2- column in Table III and the 2 row in Table IV).[14] I return to this issue later in the section.

Before turning to the Placement of the B affix, there are only a few more minor adjustments necessary to correctly generate the B forms. For reasons which seem to have little synchronic motivation, the B clitic and 'pl' are incompatible in the 1st person and 2nd person singular:

(94) *[pl +I B], *[pl +you -aug B]

These constraints can be verified by comparing the 'expected' form in the left column below with the actually occurring form to right. The expected forms contain plural -*rV*-, the occurring forms lack this affix. The corresponding form without the B affix is in the next column, showing the appearance of 'pl' -*rV* where B does not occur (next page):

[14]There is one seeming exception, namely ba-m-ba '2-.B.+' 'you (sg): them [B form].' Here the allomorph ba- is not conditioned by B, but appears for independent reasons, as is shown by the fact that it (anomalously) appears in the corresponding A-form as well: ba-ra.

(95)

		Expected B form	Occurring B form	A form	Gloss
	a.	ŋa-m-ba-*ra*	ŋa-m-ba	ŋa-[V]-*ra*	'1:+'
		1.B.+.pl	1.B.+	1.+.pl	'I: them'
	b.	na[:]-am-ba-*ra*	na[:]-am-ba	na[:]-[V]-*ra*	'1+pl: +pl'
		1+.B.+.pl	1+.B.+.	1+.+.pl	'We : them'
	c.	?ni[:]-[V]-*ri*-ŋa	na[:]-a-ŋu	ni[:]-*ri*-ŋa	'1+pl: f'
		1+.B.pl.f	1+.B.f	1+.pl.f	'We : her'
	d.	na-m-ba-*ra*	na-m-ba	na-[V]-*ra*	'12:+'
		1.B.+.pl	1.B.+	1.+.pl	'We 2: them'
	e.	ŋa[:]-am-ba-*ra*	ŋa[:]-am-ba	ŋa[:]-[V]-*ra*	'12+pl: +pl'
		1+.B.+.pl	1+.B.+.	1+.+.pl	'We all: them'
	f.	?ŋi[:]-[V]-*ri*-ŋa	ŋa[:]-a-ŋu	ŋi[:]-*ri*-ŋa	'12+pl: f'
		1+.B.pl.f	1+.B.f	1+.pl.f	'We all: her'

Cases (95a,b,d,e) might plausibly be analyzed as mere final truncations of -ra 'pl', were it not that 'pl' appears in the same circumstances in the 2nd and 3rd persons. To *ŋa-m-ba-ra, *na:-m-ba-ra, *na-m-ba-ra and *ŋa:-m-ba-ra in the 1st person (95a,b,d,e), compare the well-formed na-m-ba-[V]-ra '2.B.+.(+).pl' 'you all:them' and wam-ba-[V]-ra 'B.+.(+).pl' 'they:them.' Whereas a sequence B.+.pl is apparently ill-formed, a sequence B.+.+.pl is not. At this point, these seem to be mere accidents of the morphology. As for (95c,f), these forms, if generated with 'pl,' would surface as homophonous with the A-forms in virtue of vowel-shortening. This factor may have originally motivated the loss of plural in the 1st person B environments.[15] Later, this Impoverishment was perhaps generalized to the forms (95a,b,d,e) which even when unimpoverished, do not lead to surface homophony.[16]

We consider now the Placement of the B affix. An inspection of Table III will show that in inverse forms, the B affix appears after INV [N] when Inverse follows a person affix. However, when INV follows a gender or number affix, the B affix precedes the gender or number affix (A and M rows of Table IV):

[15] Heath gives a different explanation, citing the fact that the final /n/ of the B affix abuts the initial /r/ of the plural affix, giving a dispreferred phonological sequence.

[16] The filters in (94) constrain the operation of the Plural Enhancement rule, which, unless blocked from applying, enhances [+aug] categories by the class attribute [pl]. The failure of [pl] to appear in the forms in (95) is part of a more general trend of suppressing [pl] in complex clitic sequences, as was discussed in section 3.2.2.3.

(96) a. ŋa-*m-ba*-ŋi < ŋa-wan-ba-ŋi
1.INV.B.f 'f: 1-' 'she : me' [B form]

b. *wa*-ŋi-*ŋ*-gu
B.f.INV.A 'A: f' 'it (ANA gender) : her' [B form]

c. *wam*-ba-*ŋ*-gu
B.+.INV.A 'A: +' 'it (ANA gender): them' [B form]

In (96a), B follows INV which follows ŋa- '1.' In (96b,c), B occupies the initial position, since there is no person marker. INV sits in the second-to-last position and is separated from B by 'f' (96b) or '+' (96c).

Consider first the cases of type (96a) where INV and B come to reside between the same two affixes. Consider the derivation of such a form. As argued in the last section, all INV forms result from the Merger operation:

(97) [ŋi] * [ŋa] --> [ŋa+N+ŋi]

Now crucially, the INV [N] must be parsed as part of the person prefix in order for the B affix to be situated correctly after INV:

(98) wan * [ŋa-N- ŋi] --> [ŋa-N-wan-ŋi]

The B affixes merges as a suffix to any adjacent person affix plus INV, but remain *in situ* if followed by any other affix, as in (95b). In (98), the following person affix is the head of the following constituent because [ŋa-N- ŋi] is the head of the clitic cluster (since ŋi was the head and target of Inverse merger (97)). Initial person clitics which are not the structural heads of the clitic cluster constructions are also targets for merger, as in (99):

(99) ni-m-bi-ri-ŋa
2.B.+.pl.f '2+pl : f' 'you (pl): her [B form]'

In the above form, the B affix -m- must suffix to the initial person affix ni- even though ni- is not the head of the clitic sequence. The head of the clitic sequence is ŋa 'f,' the last clitic that is attached. At this level, then, the head of a constituent must be defined only in terms of adjacency relations, and not in terms of structural headedness. This is consistent with Marantz's (1988) treatment of second-position sentential clitics, which affix merely to initial constituent to which they are adjacent, but not to the head of the complement clause.

Thus, the INV and B are alike in that the placement of both is limited to 'second-position.' As seen in the last section, INV is limited to the second-to-last position: at most one affix may follow INV. Likewise, B is limited to second-to-first position, wherever a person marker occupies first position. Otherwise, B is initial.

These facts together suggest that the construction of the Nunggubuyu clitic sequence is a derivational operation. At some point early in the derivation, both B and the metathesizing affix which precedes INV are peripheral clitics. Some operation, which I have ascribed to Merger, then situates B and INV within the clitic sequence:

(100) B - [X * Y] - Z --> [$_X$ X+B] * [$_Y$ Z-INV-Y]

Both operations can be seen at once in certain highly complex forms such as '1+f : +.' 'we (excl, f): them' although here the presence of INV is detectable only from the [a]-harmony induced in the preceding affixes, as discussed in the last section:

(101) wan - [na: * ŋi] - wV --> [$_X$ na:+wan] * [wV-N-ŋi] --> na -am-ba-ŋi
B 1+ f + 1+.B +.INV.f

This establishes that both operation may occur together in the same form, although in most forms only one or neither will apply.

Placing of the B-affix via Merger also provides a clean account of the allomorphy of the 2nd person [-aug] affix nun-/ba-. Recall that, the ba- allomorph appears in the B form, and typically the nun- allomorph appears in the A-form:

(102) a. nun-ŋu > nu-ŋu
2-.f 'you (sg.) > her'

b. ba-ŋu
2-B.f 'you (sg.) > her [B form]'

Where the B-allomorph ba- of '2-' appears, the independent B affix wan does not. Therefore, when B merges with '2-,' the result is a fused affix, which is suppletive with the basic '2-' form nun-. The effect is similar to the suppletion of French *de* by *du* when merged with following *le*.

Analyzing B and X-INV as second-position clitics within a clitic cluster is a more explanatory account of their behavior than an account in which their Placing is simply determined ad hoc. For it is not the case that B and INV indiscriminately appear within the sequence. Rather, in the last section it was established that the requirement that INV must end up a second-position clitic motivates the majority of the Impoverishment rules governing the otherwise inexplicable 'disappearance' of expected affixes from the clitic sequences. In addition, because B realizes or is conditioned by tense/mood features, it is expected that its syntactic source (whatever it turns out to be), should be different from that of the agreement clitics: thus B should be generated outside of the clitic cluster, perhaps as the clause Complementizer. The present analysis assumes that this is so, and wherever B appears inside the agreement clitic cluster, it has essentially 'moved' there as a second-position clitic. It would be unexpected, given these assumptions, for a tense-mood clitic of any form to appear in the very middle of an agreement clitic cluster.

3.2.5 Discussion

The preceding several subsections have been devoted to exploring (1) the neutralizations which reduce the 372 combinations of arguments in transitive agreement to a mere 130 or so distinct clitic sequences in Nunggubuyu and (2) the means by which these clitics are linearly ordered. I proposed a set of filters which entail the Impoverishment of the clitic cluster at the featural level, with two results. First, certain category distinctions disappear entirely within transitive agreement. For example, object arguments which are functionally dual, trial, or plural neutralize to a single category. Second, certain 'expected' affixes fail to appear in the clitic sequence. For example, with a 3 f sg object, a 3rd person singular subject will not show gender, although gender is not *in general* neutralized for 3rd person singular subjects.

Logically, the simplest system would simply concatenate agreement affixes referring to both arguments without any neutralizations. Nunggubuyu provides very rich evidence that the logically simplest system is not the 'unmarked' state-of-affairs in natural language. As originally observed by Hale (1973), neutralization of complex categories is the norm rather than the exception in transitive agreement, *even when* transparent morphological means exist for expressing such complex categories with a ll set of affixes. This evidence alone suggests that neutralizations of various kinds are automatic and must in effect be unlearned in those systems which permit greatest complexity.

Of the Impoverishments identified for Nunggubuyu, at least half are motivated by a surface constraint on entire clitic sequences that requires them to appear in the order person > number > gender > class. Because the subject and object clitics are constructed each on their own cycle, the surface order of the whole clitic sequence may be violated when the subject and object sequences are subsequently concatenated. To meet the requirements of the surface constraint, I showed that Merger applies to repair offending sequences, and that the Merger operation yields the forms identified as Inverse by Heath. Because Merger must obey adjacency requirements in the sense of Marantz (1988), it is limited to apply only to peripheral constituents giving the well-known second-position phenomenon. I argued that Merger is the only repair operation available for reordering clitic sequences; the result is that certain clitic sequences will be unrepairable, owing to constraints on Merger.

I then showed that the unrepairable sequences are precisely those sequences which cannot be generated because of the prior application of Impoverishments. In other words, if Impoverishment does not apply to certain input representations, the derivation of a well-formed clitic sequence will *crash* in the sense of Chomsky (1992), that is to say, it will not result in a mapping to a well-formed representation at the next level. Thus, these Impoverishments may be seen as applying freely to morphosyntactic representations in order to ensure a well-formed derivation.

Merger was then argued to be the means by which the B affix arrives at the second-position in clitic sequences beginning with a person affix. I suggested that the B affix, inasmuch as it is conditioned by Tense-Mood-Aspect-Negativity, is in fact a complementizer, morphologically incorporated onto the initial clitic sequence.

The richness and internal complexity of the Nunggubuyu clitic system establishes the generative nature of the Morphological component. Knowledge of the clitic morphology of Nunggubuyu consists not merely in knowing a set of affixes and their distribution. Rather, I have shown that the clitic sequences must be generated in a derivation consisting of several steps. First, input morphosyntactic representations are Impoverished, yielding both categorial neutralizations and preventing the generation of unrepairable sequences. Second, the subject and object clitic sequences are positioned each on their own cycle. Third, the generated string is repaired by Merger when necessary. Fourth, the B affix, if present, attaches to the string, merging as a suffix to an initial person affix if present.

3.3 Participant Neutralizations in Kiowa-Tanoan and Iraqw

Of the eight Impoverishment filters proposed for Nunggubuyu, four were shown to be motivated to prevent the generation of unrepairable sequences. Others, however, apply irrespective of this consideration and therefore must be either (1) historical accidents or (2) automatically assumed filters. In this section, I show that the filter 1-2 Number deletion (or a variant thereof) applies as well in Kiowa-Tanoan and therefore is likely to be a UG filter. I then show that neutralization of subject features when the object is [+participant] occurs in Kiowa-Tanoan and also in the S. Cushitic language Iraqw.

1-2 Number Deletion for Nunggubuyu is repeated below:

(103) 1-2 Number Deletion

*[+I +aug] [+you] (hence +aug --> Ø)

The effect of this filter is that whenever both arguments are [+participant], that is, when 1 acts on 2 or vice versa, there are no number distinctions for the 1st person argument. Thus, in Nunggubuyu, the 1:2 and 2:1 combinations reduce to the following (A forms only):

(104)

		Object		
		1	2-	2+
	1	*	ŋu-nu	ŋa-na
Subject	2-	nVmba	*	*
	2+	ŋi-ri	*	*

There is no 'logical' reason why 1+ subject or object affixes could not appear in combination with '2' affixes. The following clitic sequences are nevertheless ill-formed:

(105) a. *nu :- nu
1+.2- 'we (excl.): you (sg.)'

b. *na :-na
1+.2 'we (excl.): you (pl.)'

It might be argued, for example, that (105a) is disfavored since it ends up being homophonous with the '2+: m' form; but this argument cannot apply to ill-formed (105b) which is not homophonous with any well-formed clitic sequence. Furthermore, both sequences in (105) obey the clitic sequencing constraints; the operative filter (103) cannot be motivated to prevent ill-sequenced strings.

Why then should such forms be systematically prohibited? The present analysis requires that the filter (103) be part of UG, regularly and costlessly assumed unless evidence leads the child to suppress it. It is then predicted that the same filter should appear in other unrelated morphological systems. This prediction is correct.

The Kiowa-Tanoan languages have verbal agreement with up to three arguments. When both arguments are [+participant], all Kiowa-Tanoan languages show neutralization of number similar to that seen in Nunggubuyu.

First, consider the N. Tiwa (Taos) analogues of (104) from Trager (1946):

(106)

			Object		
		1	2 sg	2 dual	2 pl
Subject	1	*	ą-	mą-pęn	mą-pi
	2	may-	*	*	*

As in Nunggubuyu, number is not distinguished for 1st person arguments when both arguments are [+participant]. Indeed the Taos system is even more restrictive than the Nunggubuyu system: when 2 acts on 1 there is but one prefix may-, which distinguishes number for neither participant. Precisely the same distribution of categories also occurs in Jemez (K. Hale, p.c.; Sprott 1992):

(107)

			Object		
		1	2 sg	2 dual	2 pl
Subject	1	*	a	mąl-	ba-
	2	bæ-	*	*	*

The S. Tiwa system (Allen, Frantz, Gardiner & Perlmutter 1990), differs minimally from that of N.Tiwa. In S. Tiwa, the form bey-, obviously cognate with N. Tiwa may- and Jem. bæ-, indicates only 2 : 1sg. For 2 : 1pl, the same prefix is used as for 2sg : 3 INV, namely ku-:[17]

(108)

				Object		
		1	1pl	2 sg	2 dual	2 pl
Subject	1	*	*	i-	men-	ma-
	2	bey-	ku-	*	*	*

[17]Whether this fact results from a pseudo-reflexive interpretation or whether 1 pl has conflated with 3 INV, as is regular in Kiowa, remains unclear to me.

Thus, the filter for S. Tiwa deletes number from the subject but not the object in a clause with two [+participant] arguments:

(109) *[$_{\text{SUBJ}}$ +part +aug] [+part]

Finally, the most radical of neutralizations occurs in Arizona Tewa (Kroskrity 1985). Where both arguments are [+participant], the passive is obligatory. 2:1 merges with 3:1 with an all-purpose 1-prefix di, but an underlying 1:2 relation has its own special prefix wí which marks the person of neither subject nor object:

(110) dí 1 subj. of a passive
wí 2, with 1-instrumental

The following sentences exemplify the use of these prefixes (Kroskrity (1985:312):

(111) a. ʔµ-di na:/na:ʔin dí-k'ú:wá-mɛgi
you-OBL I/we 1-sheep-give

'I/We was/were given sheep by you.'

b. ʔµ-n na:-di wí-tay.
you-PL I-OBL 2(1)-know

'You (nonsg) are known (or recognized) by me.'

Observe that the chômeur argument is marked by the OBL suffix -di. It is also possible, although I will not attempt to justify the matter here, that this is in fact an ergative suffix and that these clauses are ergatives instead of passives.

Arizona Tewa thus has the most restrictive filter of all the Kiowa-Tanoan languages:

(112) *[+participant +participant +aug]

As Kroskrity points out (1985:312), the Rio Grande dialect of Tewa does not neutralize number in the 1:2 forms. The relevant categories are given below from Speirs (1966):

(113)

		Subject 1	2
	1sg	*	wi
Ergative/	1du	*	wæ̂n
Chômeur	pl	*	wê
	2	dí	*

Thus, Arizona Tewa differs from Rio Grande Tewa in that Arizona Tewa has a more restrictive filter, making it impossible to generate the forms wæ̂n and wê, which then were lost from the lexicon of affixes/rules in Arizona Tewa, leaving only wi for the underlying 1:2 relation.

Additionally, it is probably not insignificant that the number distinctions which are retained in (113) are number associated with the logical subject. We saw in the Tiwa and Jemez paradigms that number distinctions of the [+participant] subject are utterly lost, and what number distinctions remain are only of the object. The Tewa prefixes however are correlated with the passive (or ergative) construction, which will essentially place the underlying 'logical' subject in an instrumental or oblique phrase. Hence, in Rio Grande Tewa, those number distinctions which are retained are those of the underlying subject and surface ergative, namely number of the 1st person category. This is consistent with the hypothesis that in an ergative-absolutive system the ergative will be 'lower' (or more marked, in terms of case) and thus will retain more morphosyntactic distinctions, just as the object will in a nominative-accusative system.

Where at least one argument is [-participant], number neutralization is considerably less extensive in all Kiowa-Tanoan languages. For example, consider the Jemez forms when [+participant] subject acts on (unaugmented) basic or (nondual) INV [-participant]. (Form given in parentheses were recorded by K. Hale and apparently reflect conservative dialects/idiolects):

(114)

	3 Object basic	INV
1	ta	te
1 du	hą (sa)	hą-pa (są-pɨ)
1 pl	he (se)	he-pe (se-pa)
2	a	e
2 du	mą (ma)	mą-pe
2 pl	ba	ba-pe

Here number remains distinct for both 1 and 2, whose three categories of number cross-classify with the various 3-Object categories.

Viewed at the level of paradigm-slots only, the Kiowa-Tanoan system is remarkably similar to Nunggubuyu. However, parsing of the actual prefixes yields a further complication. Whereas in the Nunggubuyu forms both '1' and '2' affixes are still sometimes parsable, this is never the case in Kiowa-Tanoan. In particular, the 1: 2 forms are either identical to or minimally different from the intransitive prefixes for 2nd person. For example, Jemez a, mpl, and ba mark either 1: 2 sg, du, pl respectively with an active transitive verb stem *or* 2 sg, du, pl absolutive subject with an intransitive stem. The featural identity of these pieces is merely that given below:

(115)	mpl-	2 -sg -aug	(dual)
	ba-	2 +aug	
	a-	2	

In Jemez, there seems to be filter of the following sort:

(116) $*[_{\text{SUBJ}}\ F]\ [_{\text{OBJ}}\ 2]$

In other words, with a 2 object, all the features of the subject are Impoverished, and the form surfaces with the same prefix as occurs with an absolutive 2 subject.

The filter in (116) must be kept distinct from a separate morphosyntactic phenomenon common to all Kiowa-Tanoan, namely the requirement of obligatory passive/ergative when, among other circumstances, the [-participant] acts on [+participant] (Rosen 1990 gives an interesting analysis of the circumstances in which passive is required in S. Tiwa). For example, a 3:2 clause must be expressed by a passive verb stem with 2 absolutive prefix with any 3 argument appearing with an instrumental suffix. In a 1: 2 clause, however, in Jemez (but not, as we have seen, in Tewa), the stem has its active form. Nevertheless, the 2 absolutive prefix occurs here as well. Thus, the appearance of absolutive prefixes is not fully correlated with the passive/ergative syntactic construction. Thus, I attribute the appearance of 2 absolutive prefixes on transitive stems in this circumstance to the Impoverishment in (116) rather than a 'deeper' syntactic process.

Because of the obligatory passive/ergative, any feature F in (116) will have to be a feature of a 1st person argument. This follows because if F is part of a 3rd person argument, passive/ergative is required, and if F is part of a 2nd person argument, reflexive is required. Thus the filter can be expressed in a maximally simple form as given.

The prefix system of Kiowa shows the Impoverishment of the agent in an even broader way. Consider the following prefixes for clauses with two [+participant] arguments (L Watkins 1984:115-16):

(117)

			Patient			
Agent		1sg	1pl	2sg	2du	2pl
	1sg	*	*	èm	mɔ́	bɔ́
	1pl	*	*	gɔ̀	mɔ́	bɔ́
	2sg	ę́	dɔ̀	*	*	*
	2du	mâ·	dɔ̀	*	*	*
	2pl	bâ·	dɔ̀	*	*	*

Consider first the prefixes dɔ̀, mɔ́ and bɔ́. These prefixes have the featural composition below:

(118) dɔ̀ 1 pl patient
mɔ́ 2 dual patient
bɔ́ 2 pl patient

These forms can be further decomposed, e.g. nasality marks dual, and clearly ɔ is segmentable (cf. L. Watkins 1984:117ff.), but this does not effect the general argument. The prefixes in (118) are used not only when the agent is [+participant], but also when the agent is 3rd person. (cf. Tewa dí). Thus, Kiowa has the following filter:

(119) *[AGENT F] [PATIENT +part -sg]

The singular object categories continue to distinguish number in (118), but somewhat indirectly. The prefix gɔ̀ marks any 2 sg patient, except for when the subject is 1sg, in which case the 2 sg intransitive prefix èm is used, which is, curiously, also used for 1sg:3sg clauses.

However the remaining forms in (118) are to be analyzed, the general observation is that whenever both arguments are [+participant], neutralizations of varying kinds always occur. These may have the effect of (1) erasing number distinctions, as in

Nunggubuyu or the 2 : 1 forms Jem. bæ-, Taos may-, S. Tiwa bey-, or (2) erasing all the agent features, as in Jemez ɑ-, mɒl-, bɑ-, Tewa di̦, or Kiowa dɔ̀, mɔ̀ and bɔ̀.

Similar effects can be discerned in the selector agreement system in Iraqw (Nordbustad 1988). Iraqw is a Cushitic language (Afro-Asiatic) spoken in Northern Tanzania. Selectors are preverbal particles which serve a variety of functions in the Cushitic language family. Historically, they are clearly complementizers and I will assume tentatively that this is true synchronically as well. In most Cushitic languages, for example, Arbore (Hayward 1984) or Somali, selectors vary according to the tense/mood of the clause and agree with the subject of the clause. For example, Hayward (1984:255) gives the following table, showing the agreement affix on the selector and the verb for the person-number categories in Arbore. Note that like other Afroasiatic languages discussed in chapter 1, verbal agreement splits into two positions of exponence (I have inverted the chart for easy comparison with the paradigms in ch. 1):

(120)

	selector agreement affix	verbal agreement
3m	-y	y-
3f	-y	t- -t
2s	-Ø	t- -t / t- -Ø
1s	-n̄	?-
3pl	-so	y-
2pl	-n̄	t- -t
1pl	-na(a)	n- -n

The Arbore system is reminiscent of the agreeing complementizers in Flemish and may warrant the same syntactic analysis.

Iraqw and related languages such as Gorowa, Alagwa, and Burunge (Whitely 1958) have a much richer system of selectors, which, in Iraqw at least, agree with two arguments in the clause as well as vary according to tense/mood/negativity.[18] In Iraqw, the form the selector takes is conditioned only by the object unless both arguments are 3rd person [-participant]. For example, consider the simple selector series 1 (affirmative, uncompleted action, Nordbustad 1988:222):

[18]Whitely (1958:45ff.) reports that the object selectors are less developed in Gorowa, Alagwa and Burunge, but these languages unfortunately have not been fully described as far as I am aware.

(121)

					Object					
		intr.	1	2m	2f	3m	3f	1pl	2pl	3pl
	1	a	-	u	i	u	a	-	nu	i
	2	a	i	-	-	u	a	ti	-	i
Subj.	3m	i	i	u	i	gu	ga	ti	nu	gi
	3f	i	i	u	i	gu	ga	ti	nu	gi
	3pl	i	i	u	i	ngu	ngi	ti	nu	ngi
	4pl	ta	i	u	i	ku	ka	ti	nu	ki

Observe the columns of identical forms under 1, 2m, 2f, 1pl, and 2 pl. Iraqw is like Kiowa except more extreme. In Kiowa the verbal prefix realizes no agent features when the patient is a nonsingular participant; in Iraqw, agent features never condition the shape of the selector if at least one argument is a participant.

In sum, I have provided evidence in this section for what appears to be a universal property of multiple-argument systems: the neutralization of number distinctions or of the agent features in agreement with one or more participants. Inasmuch as these neutralizations appear regularly to varying degrees in unrelated languages, this evidence supports the proposal that such neutralizations are either not very 'costly' or are automatically assumed unless specifically 'unlearned.' In general, such filters act to reduce the feature content of categories which surpass the complexity threshold of the language in question. The [+participant] categories will contribute more to overall complexity of an argument than do [-participant] categories, given that [-participant] can be underspecified in syntax (a morphosyntactically transparent feature), along the lines proposed in 0.2.2.4 for [-sg] in the presence of [dual].

3.4 The Position and Types of Verbal Inflection in Ket

This final section is devoted to a brief exposition of unusual properties in the verbal agreement system of Ket. Ket, a Yeniseyan language spoken in Central Siberia, has been the focus of intense research over the past 20 years by linguists working in (the former) Soviet Union. Nevertheless, much of the morphosyntax of Ket remains very poorly understood, although the broad outlines of an analysis are beginning to emerge. A generative analysis of Ket morphosyntax has hardly been attempted before, with Comrie

(1982) being the closest approximation; his insights provide the starting point of the present discussion.

The primary source of data for the Ket verb is a monograph devoted to the topic, Krejnovich (1968c, henceforth K.). Uspensky (1968) gives an overview of Krejnovich's data with attention to the overall structure of the verb word, and Krejnovich (1968a,b) are explorations of Aktionsart and the exponence of gender in the verb. Shabaev (1986a,b, henceforth Sa, Sb) expand on this with more careful attention to the distribution of affix types and their semantic correlates. Vall & Kanakin (1988) dispense with structural subject and object as primitives and give an analysis within a species of role and reference grammar. Theirs is the most current treatment available. These works are the main sources of data on Ket consulted here, although additional material can be found in the general grammar of Dul'zon (1968), and scores of journal and book articles appearing in the 1970s and 1980s, as well as unpublished doctoral dissertations. For works on the Ket verb, available articles include Belimov (1976, 1980a,b, 1981, 1982), Shabaev (1982), Verner (1974, 1985) and Verner & Zhivova (1981). The complex and relatively lately discovered tonal phonology of Ket is discussed in Liberman (1970) and Werner (1972, 1974, 1979). The only published lexicon, Donner (1955), is useful but sorely antiquated. Dul'zon (1972) contains a collection of texts from the several dialects of Ket.

Ket verbs inflect for up to two arguments of the clause. Some evidence suggests that these arguments are indexed according to thematic role or along the parameters of agentivity. Using data from Native American languages, Mithun (1991) has proposed a set of parameters which identify an argument as 'active' or 'agentive'; presumably similar parameters define argument roles in Ket as well. On the other hand, the relation of agreement allomorphy to semantic role is not particularly clear-cut, and may be partly a matter of lexical idiosyncrasy. The Ket system then would resemble that of Mohawk, where at least to some degree, agreement classes are determined by lexical assignment from the predicate (Baker 1989).

In some verbs, person agreement seems to be manifested twice, giving the appearance of multiple exponence of agreement properties. In such cases, I suspect that two thematic roles have been assigned: such verbs are apparently 'reflexive,' in a broad sense of the term whereby the agent and a lower thematic role-bearer are coreferent. [19]

---

[19] It is possible that in Ket the argument DPs are associated with nominals adjoined to the clause in the manner of Jelinek (1984). This is consistent with the observation of Vall & Kanakin (1988:10):

> Substantives and pronouns ... cooccur with role-markers (i.e. agreement affixes, R.N.]. They compose themselves in fairly free order, usually in pre-position to the verb and may be easily omitted, if the conditions of the communicative act do not require them [trans. mine].

Finally, the position of inflection vis-à-vis the verb stem is variable in Ket, depending on the verb stem in question, which selects the polarity of the inflectional affixes. We confine our attention here to this last consideration, since it crucially informs the parsing of the verb-word into constituents of analysis.

3.4.1 The Constituent Structure of the Verb Word

Ket verbs may begin and end with clitics. The first clitic refers to person features of the subject (or agent); the final clitic realizing plurality of this argument:

(122) person clitic - Verb - number clitic

Simple examples of such verbs can be found in unergative verbs, such as 'shake,' 'fly,' 'crawl,' 'see,' and 'listen.'

(123) a. du l'oqŋ
3 shake 'he shakes'

b. du l'oqŋ in
3 shake pl 'they shake' (Sb:142)

The position of the person and number clitics is invariant. In all the following examples, the clitics will be separated from the rest of the verb-word by a space; however, they are as far as I am aware inseparable from the verb word and within its domain phonologically. A particularly vexing aspect of Ket morphology is that proclitics consisting of a single consonant delete before following consonants (Krejnovich 1968c:25). Thus, for C-initial verbs, it is not immediately obvious whether an underlying clitic /C-/ is present. Comrie (1982) shows that this has led to unfortunate confusion in the analysis of the verbal system.[20]

Free word order alone, or even when combined with pro-drop, does not suffice to show a Jelinek-type system; other tests such as argument binding are more conclusive. (See Baker 1992 for a discussion of this topic with respect to Mohawk).

[20]The peripheral person and number clitics, distinct from AGR affixes realized more centrally on the verb word, show surprising resemblance with the Algonquian pattern as in Cree, Potowatomi, Fox and others (cf. Halle & Marantz, in progress). Three explanations for this similarity might be advanced. First, one might suppose that Ket and Algonquian descend from a common language, e.g. 'Nostratic.' This type of explanation has been fashionable among Russian scholars, but in my opinion has no merit. The second explanation contends that both the Ket and the Algonquian systems arose through parallel, but unrelated historical developments, where these particular historical developments, given a similar anterior syntax,

Within the constituent marked Verb in (122), there is considerable variation in the order of elements. Uspensky (1968) identifies 6 verb types, which may be grouped as basically *root-initial* or *root-final*: Two of these types are derivable from two others (as will be explained), so this leaves essentially four schemata:

(124) Root-Final Schemata

(1) AGR_B - T - V
(2) T- AGR_D-V

(125) Root-Initial Schemata

(1) V - AGR_B -T
(2) V- T - AGR_D

The subscripts on AGR -- B and D -- index series of agreement markers, which will be detailed. The root-final schemata are much commoner and have been the subject of most research. In some verbs, both AGR elements can cooccur, giving the order AGR_B-T-AGR_D.

The following table exhibits the various agreement markers, arranged according to the standard taxonomy, based originally on Krejnovich (1968c). There are eight series of agreement affixes: four D series and four B series. These are further subdivided into D_1, D_2, B_1, and B_2, each of which subdivisions has a further subtype A or B. Except for D_{2a} and D_{2b}, the relevance of the A/B distinction is unclear to me. Each series can also be uniquely identified by its 1(sg) and 3m(sg) members. For example, D_{1a} is also called the d/d series in the Ket literature:

were especially favored. Synchronically understood, the resemblance is an accident. An example of this type of argument is Garrett (1990) on the parallel development of NP split ergativity in Anatolian and in the Gorokan languages of Papua New Guinea, and may be summarized in a phrase from Anderson (1988): "what we find is the product not only of what can possibly exist, but also of what can come into existence." The third explanation is of a more straightforwardly theoretical nature: Ket and Algonquian, inasmuch as they have similar morphology, are exploiting similar options of a universal grammar of morphology; in particular, the syntactic mechanisms which construct complex verbs in the languages are presumed to bear some resemblance at the synchronic level.

(126) Table of Agreement affixes

	D_1 = cl		D_{2a}	D_{2b}	B_1		B_2	
	A	B	A	B	A	B	A	B
	d/d	di/du	d/a	d/Ø	ba/bu	bo/bu	ba/a	bo/o
1	d	di	d	d	ba	bo	ba	bo
2	k	ku	ku/u	ku/u	ku	ku	ku	ku
m	d	du	a/o	ø	bu	bu	a	o
f	da	də	i	ø	bu	bu	i	u
n	da	b–i	v/m	v	ø	ø	i/ø	u/ø
1 pl	d	di	daŋ	daŋ	daŋ	daŋ	daŋ	daŋ
2 pl	k	ku	kaŋ	kaŋ	kaŋ	kaŋ	kaŋ	kaŋ
3 pl	d	du	aŋ	aŋ	bu	bu	aŋ	aŋ

Gender agreement occurs only in the 3rd person singular. There are three genders which will be labelled masculine, feminine, and neuter, although the semantic correlates of gender are complex. (For a discussion in English, see Corbett 1991:19-20). Note also that *-d-* alternates with *-r-* between vowels, and other phonological processes (for example, a--> æ, k --> g --> ɣ, etc.) change these affixes in certain contexts.

Observe that for series D_1, the plural forms are identical to the respective singular forms, where 3 pl = 3 m sg. This series is the peripheral person clitic series illustrated in (123). Number distinctions are realized separately with the final peripheral number clitic, although the number clitic is sometimes omitted if there is number agreement on the verb in a D_2 series affix in addition to the initial clitic (cf. Comrie 1982).

The remaining agreement affixes shown above can be realized internal to the verb-word and form a constituent with Tense. Thus, there is a significant difference between series D_1 (peripheral) and the other series. For simplicity, I have retained the traditional classification of affixes; nevertheless it is faulty since it implicitly accords the peripheral clitic (D_1) series the same status as the remaining series (to the right of the double vertical in the table). Whether the remaining affix series are argumental clitics or AGR projections is not immediately obvious; however, given that Ket appears to have subject clitics, it would be expected for there to be object clitics as well. Henceforth, the nonperipheral agreement affixes will be labelled by series only (B, D).

There may be no internal agreement affixes at all, as in simple unergatives of the type displayed in (127) and (128). Tense has a variety of realizations: /l/ or /n/ for past, and Ø, /a/ or /s/ for nonpast.

(127) d il'-l'oqŋ
1 past-shake 'I shook' (Sb:142)

(128) d in'-doq
1 past-fly 'I flew' (K:54)

Returning to the schemata in (124) and (125), suppose that the elements in the complete verb-word except the verb root comprise a separate domain, which can be called INFL. The following generalization is then possible: within INFL, the order of elements is always: B * T * D (where * again represents strict adjacency), and INFL as a whole is either a suffix or a prefix to the verb-root:

(129) [[$_{\text{INFL}}$ B * T * D] [V]]

Other factors complicate this simple picture. The first is that quite often, a verb root will be accompanied by a stem formative which may be called a *determinative* and will be notated K (the segment /k/ is a common determinative). Within the INFL constituent, K follows B but precedes T:

(130) [$_{\text{INFL}}$ B * K * T * D]

The exact function of the determinative is unclear, however, determinatives are invariant for a given verb stem and thus seem to be part of the stem in some way.

In addition, there are a considerable number of 'extra' consonants and vowels which complicate a straightforward parsing of the verb: these are sometimes termed *interfixes* (Krejnovich 1968c:38). Common interfixes include /i, u, a ə, y, ɣ, ɢ/. These will be glossed as '0' and probably are either epenthetic (purely phonological), or mark juncture in some way. The morphological conditions in which interfixes occur have hardly begun to be understood.

Root-initial verb-words always (?) have a final 'derivational morpheme' (Krejnovich 1968c:191), which follows the INFL constituent. This will be glossed E (for stem-extension):

(131) V * [$_{INFL}$ B * K * T * D] * E

The stem extension may also be initial, giving:

(132) E * [$_{INFL}$ B * K * T * D] * V

Initial Es identified by Krejnovich (1968c:234-243) include *u- /us'-*, *e-*, *ey-/iy-*, *ik-*, *el'-*, and *al'-*. Initial E is an unproductive pattern, but final E is very common.

Examples are given below of each of the four schemata; an = sign separates the INFL constituent from the verb root and/or final E:

(133) Root-Final

a. d aŋ-t-o-l' = uŋ
3m 3pl$_B$-K-0-past-see 'he saw them' (K:65)

b. dæ q-i-l'-æ- = ril'
3f K-0-past-3f$_D$= put in 'she put herself in (e.g. a sledge)' (K:77)

(134) Root -Initial

a. d aŋ = ba-ɣ-s- i = t
3m rope =1$_B$-K-fut-0 = E 'he will bind me' (K:200)

b. d us = q-i-n-d'æŋ -i = t
3m heat = K-0-past-1pl$_D$-0 = E 'he warmed us' (K:195)

Complex verbs, formed by noun-incorporation or through derivational processes, have both pre- and post-INFL stem-pieces:

(135) a. xəgdæŋ = bo-g-i-s' = viy
snowstorm = 1$_B$-K-0-fut- wind.blow

'I will be covered up by a snowstorm, lit. I will be snowstorm-blown' (K:107)

b. sul' = ba-y = væt
sled = 1$_B$-0 = AUX 'I have a sled' (K:141fn.82).

c. dæ ik = k-i- s'-æ = siŋ
3f E = K-0-fut-0 = root 'she will tie (something) around herself' (K:242)

The [X...V], [E...V] and [V... E] forms constitute a taxonomic class referred to as discontinuous stem verbs by Ket scholars.

The initial and final clitics do not occur with a certain class of verbs whose subjects are patients or experiencers:

(136) bo-g-b-un
1-K-n-slip 'I slip'

bo-g-b-in-un
1-K-n-past-slip 'I slipped' (Sa:254)

Verbs of this type often, although not always, contain a marker *-b- < -v-*. I will gloss this as the 'n' affix (for neuter). Historically, this 'n' affix is probably the same *-v-* as the affix indexing neuter object agreement in series D_2:

(137) di v-tæt
3m 3n-hit 'I (will) hit it.' (K:62)

In general, the affix *-v-* appears in a variety of constructions, many of them with 'passive' meaning. The morphosyntactic implications of this have not yet been worked out.

The final complication is that 3rd person affixes of series D_{2a} do not follow Tense, like 1st and 2nd person D, but precede Tense.

(138) a. d il' -d-i = vit
3m past-1_{Da}-0 = make 'he made me'

b. d oŋ-o-l' = væt
3m $3pl_{Da}$-0-past = make 'he made them' (Sb:145)

Because of this idiosyncrasy, 3rd person Ds give two ordering schemata in addition to those in (124) and (125), in each case, AGR_D sits where AGR_B would.

To summarize so far, the Ket verb-word has the following structure:

(139) (cl) [(X) = [$_{INFL}$ B * K * T * D] = Y] (cl)

The constituents X and Y are filled by components of the complex verb-stem; where X is the verbal root, Y is obligatory; where Y is the verb root, X is optional.

The INFL constituent is always sandwiched between the two elements of a complex verb stem. Thus, INFL can be construed as adjacent (as suffix or as prefix) to either element of a complex verb.

One way to capture the positioning of INFL is to suppose that it merges with the left-peripheral element of a complex verb-stem at the beginning of Morphology. Merger can be illustrated as:

(140) $[\text{INFL} * [[_{V}\, X * Y]\,]] \dashrightarrow [[X + \text{INFL}]\;\; Y\,]$

When the verb stem is not complex, Merger does not take place and INFL precedes the simple verbal root.

Another possible explanation is that INFL uniformly prefixes to the *head* of verbal constituent, where head is defined structurally (and not in terms of peripherality after adjacency relations are assigned). This analysis is superior inasmuch as the behavior of INFL is constant, regardless of whether the verb is internally complex or not. For simple verbs, INFL is a prefix. This analysis predicts that the right-member of a complex verb stem is the head in all cases, since INFL must prefix to this element:

(141) $[\text{INFL}\, [\,(X)\; Y_Y\,]] \dashrightarrow [\,(X)\; \text{INFL}{+}\,Y\;_Y]$

For noun incorporation structures, where X = N, or when Y represents an Auxiliary verb which has incorporated another Verb, then Y is structurally the head and target of Merger in an obvious way:

(142) a. $[\text{INFL}\, [N\; V\;_V]] \dashrightarrow [\; N\; \text{INFL}{+}V\;_V]$
b. $[\text{INFL}\, [V\;\; \text{Aux}\;_{\text{Aux}}] \dashrightarrow [V\; \text{INFL}{+}\text{Aux}\;_{\text{Aux}}\,]$

For verbs formed with E constituents (so-called derivational morphemes), such an analysis is committed to saying that initial Es are nonheads, whereas final Es are heads. (This holds regardless of whether Es are syntactic objects or derivational affixes):

(143) a. $[\text{INFL}\, [\; E\;\; V_V]] \dashrightarrow [E\; \text{INFL}{+}V_V]$
b. $[\text{INFL}\, [\; V\;\;\; E_E] \dashrightarrow [V\; \text{INFL}{+}E_E]$

Which of these two analyses is correct must be left to future research; in particular, the exact syntactic/derivational character of the constituents of complex verbs is still unclear.

I have delayed discussion of the precise identity of the B and D agreement series within the INFL constituent. These may be actual (object) argument clitics, adjoining to the

Verb, which then raises to Tense, or they may be AGR projections. At the present state of my research, this remains an unresolved question.

3.4.2 A Preliminary Morphosyntax of the Ket INFL

I have already discussed two simple verb types: simple unergatives with peripheral clitics and no other agreement affixes (127), (128), and patient- or experiencer-subjects without peripheral clitics and with B_2 affixes (136).

(144) Agent-subject cl V
Patient subject B_2/ D_{2a} V

Other examples of the patient subject type are given below:

(145) a. ba-ɣ-æ-b = da dux
1$_{B2a}$-K-0-n = V cry 'I hear a cry' (K:205)

b. in' = æŋ - a- v = r'a
suspect = 3pl$_{B2a}$-0-n = E 'I suspect' (Sa:256)

Verbs of possession, with the possessum as an incorporated N, mark the possessor by means of a B_2 or D_{2a} affix. For B_2 as possessor, review (135b). The following examples show the use of a D_{2a} affix as possessor:

(146) a. don'-i-l'-d-i-vit
knife-0-past-1$_{D2a}$-0-have
'I knife-had'

b. don'-i-l'-gu-vit
knife-0-past-2$_{D2a}$-have
'you knife-had' (Sa:257)

In addition to the patient subject type in (145) and (146), there is another construction with one patient or experiencer argument, indexed by D_{2a} or by B_{2b}. and an expletive clitic subject dæ (or ræ after vowels), the neuter (or feminine) subject clitic of series D_{1a} (K:261ff.):

(147) expletive clitic + D_{2a}/ B_2 + V

(148) a. dæ s'ul'æj = æ-q-i-r-i =t
3n redden= 0-K- 0-1_{D2a}-0 = E
'I redden/blush'
lit., 'it reddens me'

b. dæ s'ul'æj = æ-q-i -n'-gæŋ -i =t
3n redden = 0-K 0-past- $2pl_{D2a}$-0 = E
'You (pl.) reddened/blushed'
lit., 'it reddened you (pl.)'

c. dæ s'ul'aj = bo -k-s' = a
3n redden = 1_B -K-past = E
'I redden/blush'
lit., 'it reddens me'

These verbs are always predicates where the patient argument does not control the action, hence 'blush.' They are thus unambiguously [-control] verbs in the classification of Mithun (1991). These verbs lack the *-v-/-b-* affix that sometimes appears in patient-subject verbs. The semantic difference between verbs with D and those with B is still unclear; cf. the minimal pair (148a) and (148b), where, as expected, the determinative K appears after B *-bo-* but before D *-r-*.

The presence of an expletive clitic dæ in these verb forms suggests that the peripheral clitic position is the structural subject, although why all verbs do not then require a peripheral clitic, either expletive or contentful, is an open question.

The clitic subject dæ can in fact be associated with an overt neuter nominal, as shown in (149a), but no nominal need occur, as in (149b) (K:265):

(149) a. qoɢot dæ taŋay = a-q-i-r-i = t
hunger 3n drag = 0-K-0-1_D-0 = E

'I start to suffer from hunger,'
lit. 'hunger, it drags me'

b. ul'-in aqtæ ræ taŋay = bo-l' = a
water-ABESS strongly 3n drag = 1_B-past = E

'Without water, he suffered intensely'
Lit.: 'it dragged him'

If we do not assume that the clitic dæ is the structural subject, we have no way of explaining why it occurs at all in these constructions. The facts follow straightforwardly on the assumption that the peripheral clitic is the structural subject and an (optional) overt nominal subject is adjoined to the clause as a whole.

We next consider verbs with two arguments. If both arguments are coreferent, the verb is reflexive, otherwise it is transitive. Two-place verbs have an initial subject clitic and one or more internal agreement affixes, indexing the object or object and subject. The simplest types are given in (150):

(150) a. Reflexive: cl + B_1 / D_{2b}
b. Transitive: cl + B_2/ D_{2a}

As far as I am aware,the B_1 series and the D_{2b} series are always coreferent with the structural subject; in simple transitives as in (150), this results in a reflexive interpretation, as shown in the following examples:

(151) a. dæ bu-t-i-l' = s'uk
3f $3f_{B1}$-K-0-past = push

'she pushed (herself) off (i.e. from shore in a boat)

b. k i-l'-gu = r'uk
2 0-past-2_{D2b} = move

'you (sg) moved (yourself)'

These verbs appear to manifest 'multiple exponence' of the subject agreement properties. This is a naive conception of their structure. Instead, it appears that the initial clitic is the structural subject, and the internal affix agrees with and is coreferent with the subject. B_1 and D_{2b} are thus reflexive clitics occupying the structural object position at d-structure.

Transitive verbs are similar in overall appearance to reflexives but have a different series of internal agreement affixes; either B_2 or D_{2a}:

(152) a. d i-t-i-l' = s'uk
3m $3f_{B2a}$-K-0-past = push

'he pushed her off (i.e. from the shore)' (K: 83)

b. d o-l'-gu = git
3m 0-past-2_{D2a} = look.for

'he looked for you' (K:67)

Observe that the affix series used for transitive objects are the same as those registering agreement with the single argument of unaccusative verbs. The semantic differences between B-objects and D-objects are unclear. Shabaev (1986b) suggests that a D-object is directly affected by the action of the predicate, in a situation where "the real semantic subject influences the direct object immediately, touches it, achieves in it a physical or psychological change, or causes it to act." B-objects, according to Shabaev, are affected only through the intermediary of some more directly affected object, or with verbs of affect, sensory perception, or involuntary reception of a physical state or possession. Thus, if X drags Y, then Y is a D-Object, but if X drags Y on a sled or if X sees Y, then Y is a B-Object. However, these statements reflect only loose generalizations. One possible analysis is that the D-objects and B-objects are analogous to accusative and genitive/dative objects in languages with idiosyncratic case marking, such as Georgian or Icelandic.

A more complicated class of transitive verbs show a subject clitic and two internal agreement affixes: one indexing the subject and the other the object. In the schemata below, the underlined agreement affix indexes (is coreferent with) the subject:

(153) $\underline{\mathrm{cl}}$ $\underline{\mathrm{B}}_{1}$ D_{2a}
$\underline{\mathrm{cl}}$ B_{2} $\underline{\mathrm{D}}_{2b}$

(154) a. dæ bu-g-d-i = ɢus
3f $3f_{B1}$-K-1_{D2a}-0 = carry.away

'she will carry me away' (K:93)

b. k oŋ-i-l'-gaŋ = ʁa
2 $3pl_{B2}$-0-past-$2pl_{D2b}$ = sell

'you (pl.) sold them [something]' (K:96)

It can be verified by comparing these two that the order of elements within the INFL constituent is fixed: regardless of whether B indexes the subject or the object, it is the first element of INFL, and likewise D is the last element of INFL.

This completes the different verbal types as far as I am aware. The chart below summarizes the various forms:

(155) Ket verb types

unergative	cl			
reflexive			$\underline{\text{cl}}$ $\underline{B_1}$	$\underline{\text{cl}}$ $\underline{D_{2b}}$
patient subject	B_2			
verb of possession	N...B_2	N ... D_{2a}		
expletive subject	(cl) B_2	(cl) D_{2a}		
transitive	cl B_2	cl D_{2a}	$\underline{\text{cl}}$ $\underline{B_1}$ D_{2a}	$\underline{\text{cl}}$ B_2 $\underline{D_{2b}}$

In sum, the Ket verb has an intricate constituent structure. Aside from the external subject clitic, other agreement affixes index non-agentive arguments and comprise a structural INFL unit which prefixes to the head of the complex verb. The precise details of the assignment of argument classes -- whether lexically idiosyncratic or semantically driven according to parameters of agentivity -- remain unclear but will provide an excellent source of data in future research.

3.5 Summary

In this chapter I identified three aspects whose interplay govern the realization of inflectional material in multiple-argument agreement systems. The first of these is Neutralization of categories via Impoverishment at the beginning of Morphology. The second of these is the Placing of affixes according to any of three factors: (1) the syntactic position of the M^0 which contains the features realized by the affix, (2) in cases of M^0 splitting, the hierarchy of features governs the order in which rules/affixes apply in the default cases, subject to language-specific overrides (templates), and (3) the Merger operation which relocates affixes onto the head of their complement, 'head' being defined either structurally or in terms of peripherality after adjacency relations are assigned. The third aspect is the Licensing of affixes in positions. I argued that Strict Licensing operates only when there are obligatory positions-of-exponence demanded by autonomous morphological structures (word templates). In such cases, affixes may compete for these positions; but otherwise, the failure of certain 'expected' affixes to appear was attributed to earlier feature-level neutralizations from Impoverishment.

Neutralization and Placing were shown to be partly interactive in Nunggubuyu. Of the Impoverishments which occur, four were shown to be motivated to prevent the generation of clitic sequences which could not result in well-formed strings through the available Placing mechanisms (feature-hierarchy and Merger). Impoverishment of categories when both arguments are [+participant] was shown to be a near universal occurrence from evidence in Nunggubuyu, Kiowa-Tanoan and Cushitic, and was therefore argued to be the 'unmarked' state-of-affairs in natural language. Finally, I showed that a group of affixes in the Ket verb-word comprise a structural unit (INFL), which is correctly Placed through the Merger operation.

Concluding Remarks

It is often true that the merit of a theory, linguistic or otherwise, lies not so much in the problems it solves than in the questions it brings to light. What sort of facts are now of interest, given the theory developed in this thesis?

The theory of Impoverishment, proposed in the Introduction, makes specific claims about the direction of neutralization in morphological categories. At the very least, I have proposed that there exists an unmarked direction of neutralization, obeying the hierarchy of features. Impoverishment deletes the values of certain morphosyntactic features at the beginning of Morphology, whenever a given matrix of such features surpasses a language-specific complexity threshold.

However, it should be always borne in mind that neutralization, understood as the failure of a certain expected morphosyntactic property to have a realization, may come about in ways other than Impoverishment. I have argued that neutralization occurs at several other levels. These include (1) the 'underspecification' of affixes/rules which discharge features of the input, (2) the disjunction of affixes in virtue of competition for a given obligatory position-of-exponence, and (3) the (optional) absence of redundant (enhancing) values such as [-sg] in the presence of [dual]. Because of the interplay of these factors, neutralization within morphological systems has a striking complexity. Once these four possibilities are fully recognized, each may be studied independently.

I argued that Impoverishments are not learned in most cases, but rather 'unlearned': categorial neutralizations will be assumed unless positive evidence prompts the learner to suppress them. This makes substantive claims about language acquisition which may be tested for languages with highly complex morphosyntactic categories.

The Feature Hierarchy Hypothesis and related hypotheses (Spell-Out Ordering, Impoverishment) state that there is a hierarchy of features which *in the unmarked instance* governs the direction of neutralization, affix ordering where not determined by syntax, and morphological rule ordering where not determined by Panini's principle. This too makes substantive claims about the expectations of the learner which may be tested, both through study of acquisition and through assessments of historical change.

This thesis was written in the hope that the theory of syntactic affixation could be reconciled with well-known problems in the view that each X^0 is an affix and with the view that the syntax of affixes is the syntax of X^0s. My conclusion has been that this is possible given a sufficiently articulated intermediate component, Morphology, governing the mapping between the output of syntax and the input to phonological rules. This component contains much that is idiosyncratic and language-specific; a theory of

Morphology is accordingly a theory of markedness or complexity which evaluates this idiosyncrasy as more or less costly.

It has also been my hope to draw attention to the rule/affix distinction. The current proposal recognizes both as integral to word-formation, and draws the line cleanly between them: rules change phonological information, whereas affixes add phonological information. To this rather obvious division, I then added the proviso that rules do not discharge morphosyntactic features; hence a rule is never a feature's primary exponent. What counts as a primary exponent is deducible in terms of the bleeding relations which hold among potential exponents, through nestedness of affixes as predicted by syntactic derivation, and, more indirectly, through study of historical change.

Morphologists have been greatly exercised over the past fifty years over the identity of the structuralist unit 'morpheme.' Some have abandoned it entirely, e.g. Matthews (1972a), Anderson (1981, 1992). Others have equated morphemes with morphs (Lieber 1980, Williams 1981), or morphs with X^0s (Lieber 1992), or X^0s or their analogues (Qs) with collections of morphs (Lumsden 1987, Halle 1992). I have suggested that each of these different stances has merit but fails insofar as each does not fully recognize the independence of the three types of information available in the Morphology mapping: features, positions, and affixes. On the proposal here, the units of idiosyncratic or stored information are rule-affixes, algorithmic processes which *discharge* features from syntax as phonological information, sometimes at particular positions of exponence defined by autonomous conditions on the well-formedness of words, other times 'freely' at the position of the source X^0 or in multiple, freely-split positions. Various ways were explored to encode marked mappings, such as, no isomorphy between X^0s and affixes (fusion, splitting), obligatory or limited positions-of-exponence (autonomous morphological structure), and constraints on affix ordering (the principles which derive 'templates,' and the process of Merger at the level of adjacency relations). While recognizing that such marked mappings exist, the most plausible theory of morphology will evaluate these in terms of expectations held by the language learner. It remains a clearly empirical matter how costly these deviations from a simple X^0 = affix relation are.

It has been often said that morphology is yesterday's syntax. Much of it is the unprincipled debris of the past. But not all: the intricate interplay among features, positions and affixes in Arabic and Berber, in Nunggubuyu and Ket, and in many other languages, should dispel the idea that morphology has no coherent structure unto itself. Instead, the form of words in complex inflectional systems has much to contribute to mental models of linguistic knowledge.

Bibliography

Abboud, P.F. and E.N. McCarus (1983). *Elementary Modern Standard Arabic: Part I.* Cambridge Univ. Press, Cambridge.

Abdel-Massih, E.T. (1971). *A Reference Grammar of Tamzight (Middle Atlas Berber).* Center for Near Eastern and North African Studies, Univ. of Michigan, Ann Arbor.

Aboul-Fetouh, H.M. (1969). *A Morphological Study of Egyptian Colloquial Arabic.* Mouton, the Hague.

Al-Tajir, M. A. (1982). *Language and linguistic origins in Baḥrain.* Kegan Paul International, London.

Allen, B., D. Frantz, D. Gardiner, and D. Perlmutter (1990). "Verb agreement, possessor ascension, and multistratal representation in Southern Tiwa." *Studies in Relational Grammar 3*, ed. B. Joseph and P. Postal. University of Chicago Press, Chicago.

Anderson, S. (1981). "Where's Morphology?" In *Inflectional Morphology: Introduction to the Extended Word and Paradigm Theory*, ed. T. Thomas-Flinders. Occasional Papers in Linguistics 4, UCLA. Reprinted in *Linguistic Inquiry* 13, pp. 571-612.

Anderson, S. (1986). "Disjunctive Ordering in Inflectional Morphology." *Natural Language and Linguistic Theory* 4, pp. 1-31.

Anderson, S. (1988). "Morphological change." In *Linguistics: The Cambridge Survey, vol I. Linguistic Theory: Foundations*, ed. F.J. Newmeyer. Cambridge University Press, Cambridge.

Anderson, S. (1992). *A-Morphous Morphology.* Cambridge University Press.

Anderson, S. & E.L. Keenan (1985). "Deixis." In *Language Typology and syntactic description*, v. III: *Grammatical categories and the lexicon.* Cambridge University Press, Cambridge, pp. 259-307.

Applegate, J. R. (1958). *An Outline of the Structure of* Shilḥa. American Council of Learned Societies, New York.

Archangeli, D. (1984 [1988]). *Underspecification in Yawlemani Phonology and Morphology*, MIT Ph.D diss., published by Garland Press, New York.

Arlotto, A. (1972). *Introduction to Historical Linguistics.* University Press of America, Lanham.

Aronoff, M. (1976). *Word-Formation in Generative Grammar.* MIT Press, Cambridge.

Aronoff, M. (1988). Review of *On the definition of word* by A.M. DiSciullo and E. Williams. *Language* 64:4, pp. 766-770.

Aronoff, M. (1992). "Gender in Arapesh." In *Yearbook of Morphology* 1992, ed. G. Booij and J. van Marle, Kluwer, Dordrecht.

Aronoff, M. (in progress). *Morphology by Itself*, ms. The University at Stony Brook, New York.

Babby, L. (1987). "Case. Prequantifiers, and Discontinuous Agreement in Russian." *Natural Language and Linguistic Theory* 5: 91-138.

Baker, M. (1985). "The Mirror Principle and Morphosyntactic Explanation." *Linguistic Inquiry* 16, pp. 373-416.

Baker, M. (1988). *Incorporation: a Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.

Baker, M. (1989). "Non-standard agreement in Mohawk." Paper presented at the NSF Non-Standard Case Conference, University of Arizona, July 1989.

Baker, M. (1992). "Some Subject/Object Non-Asymmetries in Mohawk and their Theoretical Impications." *Natural Language and Linguistic Theory* 10.

Basset, A. (1952). *La langue berbère*. Oxford University Press, Oxford.

Belimov, E.I. (1976). "Voprositel'nye predlozhenija v jeniseyskikh jazykakh." [Interrogative sentences in Yeniseyan]. In *Issledovanija po jazykam Sibiri*, ed. E. I. Ubrjatova. Akademia Nauk, Novosibirsk, pp. 17-27.

Belimov, E. I. (1980a). "Optativnye predlozhenija v jeniseyskikh jazykakh." [Optative sentences in Yeniseyan]. In *Podchinenie v polipredikativnykh konstrukcijakh*. ed. M.I. Cheremisina. Akademia Nauk, Novosibirsk. pp. 190-96.

Belimov, E. I. (1980b). "Otnoshenija odnorodnosti v jeniseyskikh jazykakh." [Relations of homogeneity in Yenisseyan]. In *Narody i jazyki Sibiri*, ed. E. I. Ubrjatova, Akademia Nauk, Novosibirsk, pp. 35-46.

Belimov, E. I. (1981). "Sposoby vyrazheija minimal'nogo i oslozhnennogo variantov prostogo glagol'nogo skazuemogo v jeniseyskikh jazykakh." [Means of expressing the minimal and complex variants of the simple verbal predicate in Yeniseyan.] In *Suntaksis altayskikh i evropejskikh jazykov*, ed. M.I. Cheremisina. Akademia Nauk, Novosibirsk. pp. 37-50.

Belimov, E. I. (1982). "Slozhnoe i sostavnoe skazumoe v jeniseyskikh jazykakh." [Complex and compound predicates in Yeniseyan.] In *Grammaticheskie Issledovanija po jazykam Sibiri*, ed. E. I. Ubrjatova. Akademia Nauk, Novosibirsk, pp. 132-141.

Bellegarde, J.L. & S. Ratt. (1989). *Cree, Language of the Plains*. Saskatchewan Indian Federated College.

Benveniste, E. (1935). *Origines de la formation des noms en indo-européen*. Paris, Adrien-Maisonneuve.

Benveniste, E. (1956/1971). "La nature des pronoms," in *For Roman Jakobson*, M. Halle, H.G. Lunt, H. McLean, and C. H. van Schoonevelt, eds. Mouton, The Hague, pp. 34-37. Reprinted as "The Nature of Pronouns," in *Problems of General Linguistics*, M.E. Meek, trans. University of Miami Press, Coral Gables, Florida, pp. 217-222.

Bergsträsser, G. (1928) *Einführung in die semitischen Sprachen.* Max Hueber Verlag, München. In English as *Introduction to the Semitic Languages: Text Specimens and Grammatical Sketches*, trans. P. T. Daniels. Eisenbrauns, Winona Lake, Indiana.

Bittner, M. (1918). "Charakteristik der Sprache der Insel Soqotra." *Akademie der Wissenschaft, Wien. Anzeiger. Philosophisch-historische Klasse.* Jahrg. 55: 48-83.

Bonet, E. (1991). *Morphology after Syntax: Pronominal Clitics in Romance.* PhD diss., MIT.

Brockelmann, C. (1961). *Grundriß der Vergleichende Grammatik der Semitischen Sprache. I.Laut- und Formenlehre.* Georg Olms Verlagsbuchhandlung, Hildesheim.

Buechel, E. (1970). *A Dictionary of the Teton Sioux Language*, ed. P. Manhart. Institue of Indian Studies, University of South Dakota, Vermillion, South Dakota.

Bybee, J.L. (1985). *Morphology: A Study of the Relation between Form and Meaning.* John Benjamins, Amsterdam.

Calabrese, A. (1988). *Towards a Theory of Phonological Alphabets.* PhD diss., MIT.

Calabrese, A. (1992). "The Notion of Phonological Complexity in Phonological Theory," ms. Harvard University.

Carstairs, A. (1987). *Allomorphy in Inflexion.* Croom Helm, London.

Carstairs-McCarthy (1992). *Current Morphology.* Routledge, London.

Capell, A. & H.E. Hinch (1970). *Maung Grammar, Texts and Vocabulary.* Mouton, The Hague.

Carstens, V.M. (1991). *The Morphology and Syntax of Determiner Phrases in Kiswahili.* PhD. diss., UCLA.

Cassidy, F.G. and R.N. Ringler, eds. (1971). *Bright's Old English Grammar* (Third Edition). Holt, Rinehart and Winston, New York.

Chomsky, N. (1957). *Syntactic Structures.* Mouton, The Hague.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* MIT Press, Cambridge.

Chomsky, N. (1981). *Lectures on Government and Binding.* Foris, Dordrecht.

Chomsky, N. (1989). "Some Notes on Economy of Derivation and Representation." In *MIT Working Papers in Linguistics 10*, ed. I. Laka and A. Mahajan, MIT, Cambridge.

Chomsky, N. (1992). "A Minimalist Program for Linguistic Theory." MIT Occasional Papers in Linguistics, Number 1. Distributed by MITWPL.

Clements, G.N. (1985). "The Geometry of Phonological Features." *Phonology Yearbook* 2, Cambridge Univ. Press, Cambridge.

Coate, H.H.J. & L. Oates (1970). *A Grammar of Ngarinjin, Western Australia.* Occasional Papers in Aboriginal Studies, no. 25, Australian Institute of Aboriginal Studies, Canberra.

Cohen, D. (1975). *Le parler arabe des juifs de Tunis. Tome II, Étude linguistique.* Mouton, the Hague.

Comrie, B. (1982). "Verb Agreement in Ket." *Folia Slavica* 5, pp. 115-127.

Conklin, H.C. (1962). "Lexicographic treatment of folk taxonomies." In *Problems in Lexicography*, F.W. Householder and S. Saporta, eds. Indiana University Research Center in Anthropology, Folklore and Linguistics, Bloomington, Indiana.

Corbett, G. (1991). *Gender.* Cambridge University Press, Cambridge.

Di Sciullo, A.M. & E. Williams (1987). *On the Definition of Word.* MIT Press, Cambridge.

Dixon, R.M.W. (1980). *The languages of Australia.* Cambridge University Press, Cambridge.

Donner, K. (1955). *Ketica: Materialen aus dem Ketschen oder Jenisseiostjakischen.* Compiled by A. J. Joki. Mémoires de la Société Finno-Ougrienne, 108. Helsinki.

Dul'zon, A.P. (1968). *Ketskii Jazyk.* Tomsk University Press, Tomsk.

Dul'zon, A.P. (1972). *Skazki Narodov Sibirskogo Severa I* [Tales of the Peoples of the Siberian North]. Tomsk University Press, Tomsk.

Eimas, P.D., E.R. Siqueland, P.W. Jusczyk and J. Vigorito (1971). "Speech perception in infants." *Science* 171, pp. 303-306.

Elson, B. F. (1960a). *Gramática Popoluca de la Sierra.* Instituto Lingüístico de Verano, Universidad Veracruzana, Xalapa, México.

Elson, B.F. (1960b). "Sierra Popoluca Morphology." *International Journal of American Linguistics* 26, pp. 206-33.

Elson, B.F. (1967). "Sierra Popoluca." In *Handbook of Middle American Indians.* R. Wauchope, ed., vol. 5, "Linguistics," N. A. McQuown, ed. pp. 269-290.

Emonds, J. (1976). *A Transformational Approach to English Syntax.* Academic Press, New York.

England, N.C. (1976). "The Development of the Mam Person System." *International Journal of American Linguistics* 42:3, pp. 259-261.

England, N. C. (1983). *A Grammar of Mam, a Mayan Language.* University of Texas Press, Austin.

Fabb, N. (1984). *Syntactic Affixation.* PhD diss., MIT.

Fabb, N. (1988). "English Suffixation is Constrained only by Selectional Restrictions." *Natural Language and Linguistic Theory* 6:4, pp. 527-541.

Farkas, D. (1990). "Two Cases of Underspecification in Morphology." *Linguistic Inquiry* 21:4, pp. 539-550.

Fleming, H.C. and M.L. Bender (1976). "Non-Semitic Languages," in *Language in Ethiopia*, ed. M.L.Bender et. al., Oxford University Press, London, pp. 34-58.

Forchheimer, P. (1953). *The Category of Person in Language.* Walter de Gruyter, Berlin.

Ford, C.M. (1991). "Chaha-Gurage." *Journal of Afroasiatic Languages* 2:3, pp. 231-296.

Foster, M.L. and G.M. Foster (1948). *Sierra Popoluca Speech.* Smithsonian Institution Institute of Social Anthropology, publication no. 8. United States Government Printing Office, Washington, D.C.

Fulmer, S.L. (1991). "Dual-Position Affixes in Afar: An Argument for Phonologically-Driven Morphology." *WCCFL*, pp. 189-203.

Garrett, A. (1990). "The Origin of NP Split Ergativity." *Language* 66:2, pp. 261-296.

Gazdar, G., E. Klein, G. Pullum & I. Sag (1985). *Generalized Phrase Structure Grammar.* Harvard University Press, Cambridge.

Giorgi, A. and F. Pianesi (1991, to appear). "Towards a Syntax of Temporal Representations." To appear in *Probus*. Talk given at the Eastern States Conference on Linguistics, October 1991.

Gordon, C.H. (1947a). *Ugaritic Handbook.* Pontificum Institutum Biblicum, Roma.

Gordon, C.H. (1947b). "The New Amarna Tablets." *Orientalia*, n. s. 16:1.

Greenberg, J.H. (1963). "Some Universals of Grammar with Particular Refrence to the Order of Meaningful Elements." In *Universals of Lnaguage*, ed. J. H. Greenberg. MIT Press, Cambridge.

Greenberg, J.H. (1988). "The First Person Inclusive Dual as an Ambiguous Category." *Studies in Language* 12-1, pp. 1-18.

Hale, K. (1962). "Jemez and Kiowa Correspondences in Reference to Kiowa-Tanoan." *International Journal of American Linguistics* 28:1, pp. 1-5.

Hale, K. (1973). "Person Marking in Walbiri." In *A Festschrift for Morris Halle*, S. Anderson and P. Kiparsky, eds. Holt, Rinehart and Winston, New York. pp. 308-344.

Hale, K. (to appear). "The Value of Endangered Languages." Ms. MIT.

Halle, M. (1989a). "On Abstract Morphemes and their Treatment." Paper read at the Arizona Phonology Conference, March 31, 1989. ms. MIT.

Halle, M. (1989b) "The intrinsic structure of speech-sounds." Ms. MIT.

Halle, M. (1990). "An Approach to Morphology." *Proceedings of the Northeastern Linguistics Society.*

Halle, M. (1991). "The Russian Declension." Ms. MIT.

Halle, M. (1992). "The Latvian Declension." In *Yearbook of Morphology* 1991, ed. G. Booij and J. van Marle, Kluwer, Dordrecht, pp. 33-47.

Halle, M. & A. Marantz (in progress). "Distributed Morphology and the Pieces of Inflection." Unpublished manuscript, MIT.

Harrington, J.P. (1928). *Vocabulary of the Kiowa Language.*

Harris, J. (1991). "The Exponence of Gender in Spanish." *Linguistic Inquiry* 22, pp. 27-62.

Harris, J. (1992). "The form classes of Spanish substantives." In *Yearbook of Morphology* 1991, ed. G. Booij and J. van Marle. Kluwer, Dordrecht.

Hayward, D. (1984). *The Arbore Language: A First Investigation.* H. Buske, Hamburg.

Haywood, J.A. and H.M. Nahmad (1965). *A New Arabic Grammar of the Written Language.* Harvard University Press, Cambridge.

Heath J., (1980). *Nunggubuyu Myths and Ethnographic Texts.* Australian Institute of Aboriginal Studies, Canberra.

Heath, J. (1982). *Nunggubuyu Dictionary.* Australian Institute of Aboriginal Studies, Canberra.

Heath, J. (1984). *Functional Grammar of Nunggubuyu.* Australian Institute of Aboriginal Studies, Canberra.

Heath, J. (1986) "Syntactic and Lexical Aspects of Nonconfigurationality in Nunngubuyu (Australia)." *Natural Language and Linguistic Theory.* 4, pp. 375-408.

Hercus, L.A. (1966). "Some Aspects of the Form and Use of Trial Number in Victorian Languages and in Arabana." *Mankind* 6, pp. 335-37.

Herdner, A. (1938). "Une particularité grammaticale commune aux textes d'el-Amarna et de Ras-Shamra: *t-* préfixe pronominal de la troisième personne masculin pluriel." *Revue des études sémitiques* 1938, pp. 76-83.

Hockett, C.F. (1966). "What Algonquian is really like." *International Journal of American Linguistics.* 32:1, pp. 59-73.

Hodge, C.T. (1971). "Afroasiatic: an Overview." In *Afroasiatic: A Survey* ,ed. C.T. Hodge. Mouton, The Hague, pp. 9-26.

Horn, L.R. (1984). "Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature." In *Meaning, form and use in context: Linguistic applications*, ed. D. Schiffrin. Georgetown University Press, Washington, pp. 11-42.

Hudson, R.A. (1974). "A Structural Sketch of Beja." *African Language Studies* 15, pp. 111-142.

Huntley, D. (1980). "The evolution of genitive-accusative animate and personal nouns in Salivic dialects." In *Historical Morphology*, ed. J. Fisiak. Mouton, the Hague.

Hutchisson, D. (1986). "Sursurunga Pronouns and the Special Uses of Quadral Number." In U. Wiesemann, ed. *Pronominal Systems*, pp. 1-20. Gunter Narr Verlag, Tübingen.

Hyman, L. (1979). *Aghem Grammatical Sturcture*. Southern California Occasional Papers in Linguistics 7. Department of Linguistics, University of Southern California, Los Angeles.

Ingram, D. (1971/1978). "Typology of Universals of Personal Pronouns." In *Universals of Human Language*, vol. 3, J.H. Greenberg, ed. Stanford University Press, Stanford, pp. 214-247. Reprinted from *Working Papers on Language Universals* 5.

Itô, J. (1986). *Syllable Theory in Prosodic Morphology*. PhD diss. Univ. of Massachusetts at Amherst. Published 1988 by Garland Publishers, New York.

Ivanov, V. , V.N. Toporov & B.A. Uspensky, eds. (1968)*. *Ketskii Sbornik: Lingvistika* [Studia Ketica: Linguistics], M.V. Lomonosov University, Institute of Oriental Languages, Moscow, 1968.

Jakobson, R. (1942). "The Paleosiberian Languages." *American Anthropologist* 44, pp. 602-620.

Jelinek, E. (1984). "Case and Configurationality." *Natural Language and Linguistic Theory* 2, pp. 39-76.

Jelinek, E. (1992). "Ergativity and Argument Type." Lecture given at the MIT Ergativity Project, April 30, 1992.

Jensen, J.T. (1990). *Morphology: Word Structure in Generative Grammar*. John Benjamins, Amsterdam.

Jensen, J.T. and M. Stong-Jensen (1984). "Morphology is in the Lexicon!" *LI* 15, pp. 474-498.

Johnstone, T.M. (1987). *Mehri Lexicon and English-Mehri Word-List*. Compiled by G. R. Smith. School of Oriental and African Studies, London.

Jolley, C. (1983). "Algonquian Person Hierarchy: Morphosyntactic or Semantic?" *Actes du Quatorzième Congrès des Algonquianistes*.

Kaye, A.S. (1976). *Chadian and Sudanese Arabic in the light of Comparative Arabic Dialectology*. Mouton, the Hague.

Kiparsky, P. (1973). "Elsewhere in Phonology." In *A Festschrift for Morris Halle*, S. Anderson and P. Kiparsky, eds. Holt, Rinehart and Winston, New York, pp. 93-106.

Kiparsky, P. (1982). "Lexical Phonology and Morphology." In *Linguistics in the Morning Calm*, ed. I.S. Yang. Hanshin, Seoul.

Kiparsky, P. (1983). "Word formation and the lexicon." In *Proceedings of the 1982 Mid-America Linguistics Conference*, F. Ingemann, ed, University of Kansas, Lawrence, Kansas.

Klingenheben, A. (1956). "Die Präfix- und die Suffixconjugationen des Hamitosemitischen." *Mitteilungen des Instituts für Orientforschung* 4, Deutsche Akademie der Wissenschaft zu Berlin, pp. 211-77.

Krejnovich, E.A. (1968a). "Sposoby dejstvija v glagole ketskogo jazyka." [Action-types in the Ket verb.] In Ivanov, Toporov & Uspenskiy, eds., pp. 75-138.

Krejnovich, E.A. (1968b). "O grammaticheskom vyrazhenii imennykh klassov v glagole ketskogo jazyka." [On the grammatical expression of noun classes in the Ket verb.] In Ivanov, Toporov & Uspenskiy, eds., pp. 139-195.

Krejnovich, E.A. (1968c). *Glagol ketskogo jazyka.* Akademia Nauk, Leningrad.

Kroskrity, P. (1985). "A Holistic Understanding of the Arizona Tewa Passives." *Language* 61:2, pp. 306-328.

Laka, I. (1990). *Negation in Syntax: On the Nature of Functional Categories and Projections.* PhD. diss., MIT.

Lambdin, T.O. (1971). *Introduction to Biblical Hebrew.* Charles Scribner's Sons, New York.

Lambdin, T.O. (1978). *Introduction to Classical Ethiopic Geʕez.* Harvard Semitic Studies 24. Scholars Press, Missoula, Montana.

Lasnik, H. (1981). "On Two Recent Treatments of Disjoint Reference." *Journal of Linguistic Research* 1: 4.

Li, Y. (1990). *Conditions on X^0 Movement.* PhD diss., MIT.

Liberman, A.S. (1970). "Slogovaja akcentuacija ketskogo jazyka." [Ket syllable accentuation.] In *Lingvisticheskije issledovanija* 1, Leningrad, pp. 371-389.

Lieber, R. (1980). *The Organization of the Lexicon.* PhD diss., MIT.

Lieber, R. (1987). *An Integrated Theory of Autosegmental Processes.* State University of New York Press, Albany.

Lieber, R. (1989). "On Percolation." In *Yearbook of Morphology* 2, pp. 95-138.

Lieber, R. (1992). *Deconstructing Morphology: Word Formation in Syntactic Theory.* University of Chicago Press, Chicago.

Lipkind, W. (1945). *Winnebago Grammar*, New York.

Lumsden, J. (1987). *Syntactic Features: Parametric Variation in the History of English.* MIT PhD. diss.

Lumsden, J. (1992). "Underspecification in Grammatical and Natural Gender." *Linguistic Inquiry* 23:3, pp. 469-486.

Marantz, A. (1988). "Clitics, Morphological Merger, and the Mapping to Phonological Structure." In *Theoretical Morphology: Approaches in Modern Linguistics*, ed.M. Hammond and M. Noonan. Academic Press, Orlando.

Marantz, A. (1992). "What Kind of Pieces are Inflectional Morphemes." Paper Presented at the *Berkeley Linguistics Society*, February 1992.

Marçais, Ph. (1977). *Esquisse grammaticale de l'arabe maghrebin.* Maisonneuve, Paris.

Matthews, P.H. (1972a). "Huave verb morphology: some comments from a non-tagmemic viewpoint." *IJAL* 38, pp. 96-118.

Matthews, P.H. (1972b). *Inflectional Morphology.* Cambridge University Press, Cambridge.

Matthews, P.H. (1974). *Morphology: An Introduction to the Theory of Word-Structure.* Cambridge, Cambridge University Press.

McCarthy, J. (1979). *Formal Problems in Semitic Phonology and Morphology.* PhD diss., MIT. Published 1982 by Garland Publishers, New York.

NcCarthy, J. (1992). " 'By the Rivers [in] Babylon': Akkadian Prosodic Morphology." Talk given at MIT Linguistic Colloquium, Nov. 22, 1991.

McCarthy, J. and A. Prince (1990a). "Prosodic Morphology and Templatic Morphology." In *Perspectives on Arabic Linguistic II,* Current Issues In Linguistic Theory 72, ed. M. Eid and J. McCarthy. John Benjamins, Amstersam, pp. 1-54.

McCarthy, J. and A. Prince (1990b). "Foot and Word in Prosodic Morphology: The Arabic Broken Plural." *Natural Language and Linguistic Theory* 8:2, pp. 210-284.

McKay, G. (1981). "Pronominal Person and Number Categories in Rembarrnga and Djeebbana." *Oceanic Linguistics* 17.

Meinhof, C. (1912). *Die Sprache der Hamiten.* L. Friederichsen, Hamburg.

Merrifield, W.R. (1959). "Classification of Kiowa Nouns." *International Journal of American Linguistics* 25:4, pp. 269-71.

Mithun, M. (1991). "Active/agentive case marking and its motivations." *Language* 67:3, pp. 510-546.

Mohanan, K.P. (1986). *The Theory of Lexical Phonology.* Reidel, Dordrecht.

Mohanan, K.P. (1991). "On the bases of radical underspecification." *Natural Language and Linguistic Theory* 9:285-325.

Moscati, S., ed. (1964). *An Introduction to the Comparative Grammar of the Semitic Languages: Phonology and Morphology*. Otto Harrassowitz, Wiesbaden.

Mukarovsky, H.G. (1983). "Pronouns and Prefix Conjugation in Chadic and Hamito-Semitic." In *Studies in Chadic and Afroasiatic Linguistics*, ed. E. Wolff and H. Meyer-Bahlburg. H. Buske Verlag, Hamburg, pp. 51-63.

Murtonen, A. (1967). *Early Semitic*. E.J. Brill, Leiden.

Nordbustad, F. (1988). *Iraqw Grammar*. Dietrich Reimer, Berlin.

Noyer, R. (1991). "The Mirror Principle Revisited: Verbal Morphology in Maung and Nunggubuyu." in *MIT Working Papers in Linguistics 14*, pp. 195-209.

Ntagne, S. & G. Sop (1972). *Pé ŋKé: Manuel de Bamiléké à l'usage de 6°*. Collège Libermann, Douala, Cameroon.

Pesetsky, D. (1991). *Zero-Syntax*. Ms., MIT.

Pollock, J.-Y. (1989). "Verb Movement, Universal Grammar, and the Structure of IP." *Linguistic Inquiry* 20, pp. 365-424.

Prasse, K.- G. (1963). "Les affixes personnels du verbe berbère (touareg)." *Acta Orientalia* 27, pp. 11-21.

Prasse, K. - G. (1971). "Die dialektale Einteilung des Tuareg und ihre Kriterien." In *Afrikanische Sprachen und Kulturem -- Ein Querschnitt* [Festschrift J. Lukas], ed. V. Six et. al. Hamburger Beiträge zur Afrika-Kunde 14, Hamburg, pp. 201-208.

Prasse, K.- G. (1973). *Manuel de grammaire touarègue: VI-VII, Verbe*. Akademisk Forlag, Copenhague.

Prochazka, Th. Jr. (1988). *Saudi Arabian Dialects*. Kegan Paul International, London.

Qafisheh, H.A. (1977). *A Short Reference Grammar of Gulf Arabic*. University of Arizona Press, Tuscon.

Raz, S. (1983). *Tigre Grammar and Texts*. Undena Publications, Malibu, California.

Reiner, E. (1966). *A Linguistic Analysis of Akkadian*. Mouton, the Hague.

Rosen, C. (1990). "Rethinking Southern Tiwa: The Geometry of a Triple Agreement Language." *Language* 66:4, pp. 669-713.

Rossi, Conti (1913). "Schizzo del dialetto Saho dell'Alta Assaorta in Eritrea." *Rendiconti della Reale Accademia dei Lincei* 22, fasc. 5, pp. 151-246.

Sagey, E. (1986). *The Representation of Features and Relations in Non-Linear Phonology*. MIT Ph.D dissertation.

Schwartz, L.J. (1979). "The Order of Dakota Person Affixes: An Evaluation of Alternatives." *Glossa* 13:1, pp. 3-13.

Selkirk, E. (1982). *The Syntax of Words*. MIT Press, Cambridge.

Shabaev, V. G. (1982). "Ispol'zovanie chasticy *bin* dlja vyrazhenija vidovogo znachenija zevershennosti v ketskom jazyke." [The use of the particle *bin* in the expression of perfective aspectual meaning in Ket.] In *Grammaticheskie Issledovanija po jazykam Sibiri*, ed. E. I. Ubrjatova. Akademia Nauk, Novosibirsk, pp. 153-158.

Shabaev, V.G. (1986a). "Projavlenie chert aktivnosti i ergativnosti v ispol'zovanii lichnykh pokazateley ketskogo glagola." [Manifestations of active and ergative traits in the use of person markers in Ket.] In *Paleoaziatskie Jasyki: sbornik nauchnykh trudov*, ed. P. Ja. Skorik. Akademia Nauk, Leningrad, pp. 251-261.

Shabaev, V.G. (1986b). "Realizacii semanticheskoj determinanty 'aktivnost' - inaktivnost' v strukture predikativnoj sintagmy ketskogo jazyka." [The Realization of the semantic determinant 'actvity - inactivity' in the sturcture of the Ket verb predicateive syntagm.' In *Sintaksicheskaja i leksicheskaja semantika (na meteriale jazykov raznyx sistem)*, ed. L. T. Panin, M. I. Cheremisina, Akademia Nauk, Novosibirsk.

Siegel, D. (1977). "The Adjacency Condition and the Theory of Morphology." *NELS* 8, pp. 198-97.

Silverstein, M. (1976). "Hierarchy of features and ergativity." In *Grammatical Categories in Australian Languages*, ed. R.M.W.Dixon. Australian Institute of Aborginal Studies, Canberra, pp. 112-171.

Skorik, P. Ja. (1977). *Grammatika čukotskogo jazyka, tom II.* Moscow, Akademia Nauk.

Speas, M. J. (1990). *Phrase Sturcture in Natural Language*. Kluwer, Dordrecht.

Speirs, R.H. (1966). *Some Aspects of the Structure of Rio Grande Tewa*. PhD thesis, SUNY, Buffalo. Reproduced by University Microfilms, Ann Arbor, Michigan.

Sproat, R. (1985). *On Deriving the Lexicon.* MIT PhD dissertation, distributed by MITWPL.

Sprott, R.W. (1992). *Jemex Syntax*. PhD. diss., University of Chicago.

Stairs, E. F. and E. E. de Hollenbach (1981). *Gramática Huave*, trans. into Spanish by J. B. Magaña. Appears in *Diccionario Huave de San Mateo del Mar*, by G.A.F. Kreger & E.E.Scharfe de Stairs. Summer Institute of Linguistics, Mexico.

Steriade, D. (1987). "Redundant Values." In *Parasession on Autosegmental and Metrical Phonology*, ed. A. Bosch. *Chicago Linguistics Society* 23.

Stevens, K. N. and S. J. Keyser (1989). "Primary features and their enhancement in consonants." *Language* 65:1, pp. 81-106.

Streeter, L.A. (1976). "Language perception of two-month old infants shows effects of both innate mechanisms and experience." *Nature* 259, pp. 39-41.

Suárez, J. (1975). *Estudios huaves.* Instituto Nacional de Antropologia e Historia, Mexico.

Taraldsen, K.T. (1981). "Case Conflict in Norwegian Topicalization." *Proceedings of NELS* 11, GLSA, University of Massachusetts, Amherst.

Thackston, W.M. Jr. (1984). *An Introduction to Koranic Arabic.* Harvard University Dept. of Near Eastern Languages and Civilizations, Cambridge.

Thomas, Bertram (1937). *Four Strange Tongues from South Arabia: The Hadara Group.* Proceedings of the British Academy, Vol. XXIII. Humphrey Milford Amen House, London.

Trager, G.L. (1946). "An Outline of Taos Grammar." In *Linguistic Structures of Native America,* ed. Cornelius Osgood. Viking Fund Publications in Anthropology #6. Viking, New York, pp. 184-221.

Trager, G.L. (1954). "Taos II: Pronominal Reference." *International Journal of American Linguistics* 20:3, pp. 173-180.

Travis, L. (1984). *Parameters and the Effects of Word Order Variation.* PhD diss., MIT.

Uspenskiy, B.A. (1968). "O sisteme ketskogo glagola." [On the Ket verbal system.] In Ivanov, Toporov & Uspenskiy, eds., pp. 196-228.

Vall, M.I. & I.A. Kanakin (1988). *Kategorii glagola v ketskom jazyke.* [Verb Categories in Ket.] Akademia Nauk, Novosibirsk.

van den Branden, A. (1969). *Grammaire phénicienne.* Librairie du Liban, Beyrouth.

Verner, G.K. (1974). "Reliktovye priznaki aktivnogo stroja v ketskom jazyke." [Relic features of active structure in Ket.] *Voprosy Jazykoznanija* 1974, no. 1.

Verner, G.K. (1985). "Jeniseyskoe jazykoznanie: rezyl'taty i perspektivy." [Yenisseyan linguistics: Results and prospects.] *Voprosy Jazykoznanija* 34:3, pp. 33-42.

Verner, G.K. & G.T. Zhivova (1981). "K Kharakteristike klassnoy sistemy v jeniseyskikh jazykakh." [Towards a description of the class system in Yenisseyan.] *Voprosy Jazykoznanija* 5. pp. 31-38.

Vogt, H. (1971). *Grammaire de la langue géorgienne.* Universitetsforlaget, Oslo.

von Soden, W. (1969). *Grundriß der Akkadischen Grammatik.* Analecta Orientalia 33/47. Pontificum Institutum Biblicum, Roma.

Voorhoeve, J. (1967). "Personal Pronouns in Bamileke." *Lingua* 17, pp. 421-430.

Watkins, C. (1962). *Indo-European Origins of the Celtic Verb.* Dublin Institute for Advanced Studies, Dublin.

Watkins, L.J. (1984). *A Grammar of Kiowa.* University of Nebraska Press, Lincoln, Nebraska.

Werker, J.F. (1989). "Becoming a Native Listener." *American Scientist* 77, pp. 54-59.

Werner, H.K. (1972). "Die Tonität der gegenwärtigen Jenisseischen Dialekte." *Zetischrift für Phonetik* 25, pp. 111-125.

Werner, H.K. (1974). "Die Akzentuation der mehrsilbigen Wörter in den gegenwärtigen Jenissei-Dialekten." *Zeitschrift für Phonetik* 27: 5, pp. 379-386.

Werner, H. K. (1979). "Zur Frage dem Ursprung der Tonalität in den Jenissei-Sprachen." *Zeitschrift für Phonetik* 32, pp. 181-185.

Whiteley, W.H. (1958). *A Short Description of Item Categories in Iraqw (with material on Gorowa, Alagwa and Burunge).* East African Institute of Social Research, Kampala, Uganda.

Wiesemann, U., ed. (1986). *Pronominal Systems.* Gunter Narr Verlag, Tübingen.

Williams, E. (1981). "On the notions 'lexically related' and 'head of a word.'" *Linguistic Inquiry* 12, pp. 245-274.

Wright, W. (1951). *A Grammar of the Arabic Language*, 3rd ed. Cambridge University Press, Cambridge.

Wonderly, W., L.F. Gibson, and P.L. Kirk (1954). "Number in Kiowa: nouns, demonstratives, and adjectives." *International Journal of American Linguistics* 20, pp. 1-7.

Wurzel, W. (1989 [1984]). *Inflectional Morphology and Naturalness.* Reidel, Dordrecht. Originally published as *Flexionsmorphologie und Natürlichkeit* 1984. Akademie Verlag, Berlin.

Yip, M. (1991). "Three Approaches to Reduplication with Fixed Melodic Material: Arguments for Overwriting Association." Paper presented at the *North Eastern Linguistic Society* XXII, Oct. 27, 1991, University of Delaware.

Zaborski, A. (1974). *The verb in Cushitic.* Studies in Hamito-Semitic I. Prace Językoznawcze 48, Krakow Uniw.

Zeltner, J.-C. & H. Tourneux (1986). *L'arabe dans le bassin du Tchad.* Éditions Karthala, Paris.

Zwicky, A. (1977). "Hierarchies of Person." *Papers from the Thirteenth Meeting, Chicago Linguistics Society*, W.A.Beach, S.E. Fox, and S. Philosoph, eds., pp. 714-733.

Zwicky, A. & G. Pullum (to appear). "A Misconceived Approach to Morphology." *Proceedings of the Tenth West Coast Conference on Formal Linguistics* 1991.

ABBREVIATIONS

A	1. ergative agreement [Mam] 2. ANA noun class [Nunggubuyu]
acc	accusative
ag	aggregate [Tewa]
AG	agent
AGR	agreement
anim	animate
Anm	Anmerkung [remark]
Ar	Arabic
ARG	argument [Catalan]
ATR	advanced tongue root
aug	augmented number
aux	auxiliary verb
B	1. absolutive agreement [Maung] 2. a tense/mood/aspect/negativity affix [Nunggubuyu] 3. an object agreement clitic [Ket]
B.C.E.	Before Common Era = B.C.
C	[+consonant]
c	class feature = inanimate genders [Nunggubuyu]
CL, cl	clitic
cons	consonantal
COR	coronal
CT	control transitiver [Lummi]
D	1. Determiner 2. an agreement clitic [Ket]
Dan	Daniel
dis	distributed [Tewa]
DFL	default affix
DORS	dorsal
du	dual
E	stem extension or derivational element [Ket]
E.C.	Elsewhere Condition
ENCL, encl	enclitic [Mam]
Eng	English
env	environment
ERG	ergative
ex	example
excl	exclusive
EWP	Extended Word-Paradigm theory
F	feature
f, fem	feminine
FUT	future
g	gender
gen	genitive
Gen	Genesis
Gm	German
Heb	Hebrew
impf	imperfect
inan	inanimate
incl	inclusive
INDEF	indefinite
indic	indicative

INFL	inflection
instr	instrumental
INV	inverse
Jem	Jemez
K	1. Kiowa
	2. a determinative affix [Ket]
LAB	labial
LF	Logical Form
M	MANA class [Nunggubuyu]
M^0	syntactic constituent after rebracketing, merger and/or fusion
MASC, m	masculine
Meḥ	Meḥri
N	noun, nominal
n	1. number affix [Nunggubuyu]
	2. neuter [Ket]
nom	nominative
OBJ	object
obl, OBL	oblique
p	person affix [+participant] in Nunggubuyu
PAT	patient
PERF	perfective
PF	Phonological Form
pl	plural = [-singular]
PLUR	plural gender [Nunggubuyu]
POSS	possessive
prt	participant
ptcl	particle
Px	possessor affix
Q	obligatory position of exponence in autonomous morphological structure
refl	reflexive
s, sg	singular
Sam	Samuel
SUB	subject
T	tense
T-	t-prefix in Semitic
TH	theme
Tig	Tigre
V	verb, or vowel = [-consonantal]
X^0	atoms of syntactic representation
0	interfix (epenthetic null element) [Ket]
1	1st person = [+I]
2	2nd person = [+you]
12	1st person inclusive = [+I +you]
3	3rd person = [-participant]
-	[-augmented] [Nunggubuyu]
+	[+augmented] [Nunggubuyu]
:	separates agent from patient in glosses of verbal agreement
*	1. ill-formed, or reconstructed (before a form)
	2. strict adjacency (in strings)
	3. in glosses, inanimate [Nunggubuyu]
>	precedes (in linear order), is higher than (on the hierarchy of features)
<	derives from etymologically