

Concreteness in Grammar

The Noun Class Systems of
the Arapesh Languages

Lise M. Dobrin



Concreteness in Grammar

Stanford Studies in Morphology and the Lexicon

A series edited by

Farrell Ackerman

Beth Levin

Andrew Spencer

Arnold Zwicky

The goal of this series is to provide linguists and researchers in related areas with theoretically informed works that focus on morphology as an independent subdiscipline, as well as on the interface of morphology with the lexicon, syntax, and semantics.

Concreteness in Grammar

The Noun Class Systems of
the Arapesh Languages

Lise M. Dobrin

Copyright © 2012
CSLI Publications
Center for the Study of Language and Information
Leland Stanford Junior University
Printed in the United States
16 15 14 13 12 1 2 3 4 5

Library of Congress Cataloging-in-Publication Data

Dobrin, Lise M.
Concreteness in grammar: the noun class systems of the Arapesh languages
/ Lise M. Dobrin.

p. cm. — (Stanford studies in morphology and the lexicon)
Includes bibliographical references and index.

ISBN: 978-1-57586-607-9 (cloth : alk. paper)

ISBN: 978-1-57586-606-2 (pbk. : alk. paper)

1. Mountain Arapesh language—Noun. 2. Southern Arapesh
language—Noun. I. Title. II. Series: Stanford studies in morphology and the
lexicon.

PL6621.A7D63

2012

499'.12—dc23

2012032215

CIP

eISBN: 978-1-57586-656-7 (electronic)

*CSLI Publications gratefully acknowledges a generous gift from Jill and
Donald Knuth in support of scholarly publishing that has made the
production of this book possible.*

∞ The acid-free paper used in this book meets the minimum requirements
of the American National Standard for Information Sciences—Permanence
of Paper for Printed Library Materials, ANSI Z39.48-1984.

CSLI was founded in 1983 by researchers from Stanford University, SRI
International, and Xerox PARC to further the research and development of
integrated theories of language, information, and computation. CSLI headquarters
and CSLI Publications are located on the campus of Stanford University.

CSLI Publications reports new developments in the study of language,
information, and computation. In addition to lecture notes, our publications
include monographs, working papers, revised dissertations, and conference
proceedings. Our aim is to make new results, ideas, and approaches available as
quickly as possible.

Please visit our web site at
<http://cslipublications.stanford.edu/>
for comments on this and other titles, as well as for changes
and corrections by the author and publisher.

Contents

List of Illustrations vii

Preface ix

Symbols and Abbreviations xv

1 Introduction 1

1.0 Overview 1

1.1 The Arapesh Languages 3

1.1.1 Historical and Typological Orientation 3

1.1.2 Arapesh Varieties and Sources 5

1.2 Noun Categorization and Word Form 9

1.3 The Phonological Basis of Arapesh Noun Classification 14

1.3.1 How is Arapesh Special? 14

1.3.2 Abu' 17

2 Noun Classification in Rohwim Arapesh 27

2.0 Introduction 27

2.1 The *Morphology by Itself* Model 28

2.2 Arapesh Nouns 31

2.3 The Mapping between Phonological Form,
Morphological Realization, and Morphosyntactic Class 43

2.3.1 Aronoff's Analysis 43

2.3.2 Subverting the 'Double Inversion' 44

3	The Lexical Representation of Arapesh Nouns	59
3.0	Introduction	59
3.1	The Lexical Representation of Singulars	60
3.2	The Lexical Representation of Plurals	71
4	Organization within and across Noun Classes	81
4.0	Introduction	81
4.1	The Internal Structure of Noun Class Categories	81
4.2	Organization across Noun Class Categories	95
5	Arapesh Defaults	111
5.0	Default Categories	111
5.1	More on the Default Plural <i>-ehas</i>	112
5.1.1	One Rule or Many?	112
5.1.2	Is the <i>-ehas</i> Plural ‘Regular’?	114
5.2	Default Agreement in Arapesh	121
5.2.1	Types of Formal Noun Class Assignment	121
5.2.2	The Special Behavior of <i>s</i> -final Borrowings	126
5.2.3	The Phonological Basis of the Arapesh Classes	132
5.3	A Contrastive Case: Noun Class Restructuring in Bantu	141
6	Theoretical Consequences of Alliterative Concord	155
6.0	Introduction	155
6.1	The Relationship between Word Form and Syntax	156
6.2	Two Proposed PPFS Violations in Agreement	158
6.2.1	Somali	158
6.2.2	Vata	159
6.3	Two Apparent PPFS Violations in Local Agreement	163
6.3.1	Godié	163
6.3.2	Abu’	164
6.4	Alliterative Concord in Bainuk	168
6.5	Implications for Grammatical Organization	171
7	Conclusion	183
	References	189
	Index	205

List of Illustrations

Figure

Map of the Arapesh Language Area 8

Tables

Table 1. The Arapesh Noun Class Assignment System 138

Table 2. The Classification of Selected Arapesh Nouns 139

Preface

When I was expecting my first child, a relative gave me a Jewish baby-naming book, *The Name Dictionary: Modern English and Hebrew Names*. While the book was indeed useful to my husband and me as a source for potential baby names, what I found most remarkable about it was a polemical treatise it included on contemporary Jewish naming practices and their semiotic pitfalls. Jews today typically have both a Hebrew name and a secular vernacular name. According to the author, this duality presents a problem: ‘How shall these two names be selected? How can we make them harmonize with each other?’ (Kolatch 1982:vii). These questions, which are ultimately questions about linguistic motivation and mapping, are addressed in the book’s fascinating appendix, ‘The History and Development of Personal Names’. Here two common methods for relating Hebrew and vernacular names, the ‘translation method’ and the ‘assonance method’, are critically compared. (Both methods have apparently been in use since the third century B.C.E.) In the translation method, the vernacular name is chosen for its ability to express the same core meaning as the corresponding Hebrew name (presumed to be assigned first). The result is pairings such as Hebrew *Zadok* and Greek *Justus* ‘justice’; Hebrew *Chayim* and Old French *Vive* ‘life’; and Hebrew *Menachem* and Italian *Tranquillus* ‘comfort’. In the assonance method, the vernacular name is chosen on the basis of its similarity in sound to the Hebrew name. Often the similarity is limited to the initial portion of the word. Following the logic of assonance, for example, *Menachem* was paired with Greek *Menelaus*, *Joshua* with *Jason*, and *Elijah* with *Elas*. As it happened, we decided to name our own son *Elie* after his deceased maternal great-uncle *Alfred*. Both these names correspond, on the logic of assonance, not only to one another but to the initial sounds of the Hebrew name the two individuals also share: *Elchanan*.

But according to the author of *The Name Dictionary*, our decision was an unfortunate one, as it was driven by the same faulty logic that leads to so many other ‘far-fetched, and often silly results that we witness today in connection with naming after relatives’ (331). Due to the ‘unconscious and widespread adoption of the assonance method’ (331), many name-givers

deceive themselves into believing that they are naming their child after a relative when they select a name beginning with the first letter of the name of the deceased. John and Joseph, as well as Barton and Benjamin, are as distant as day is from night. Like day and night they meet only at the extremes, but their essence is completely different. Two names may have a first letter or last letter in common, yet remain completely foreign to one another. [Kolatch 1982:334]

Assonance is acknowledged to have certain merits inasmuch as it provides ‘an emotional tie with the past and serves as a reminder of that which went before’ (330), but these features do not do not serve to ‘harmonize’ two names unless they also share meaning. So while similarity of sound may be acceptable and even desirable when applied as a secondary principle, it alone is insufficient to relate two linguistic items and hence to properly motivate the pairing of names.

What interests me most about this book is not the practical advice it offers expecting parents (we obviously disregarded it when we named our son!), but the conviction it brings to the question of the relative value of form vs. meaning in motivating the relationship between linguistic signs. In its devaluation of form it instantiates a semiotic ideology argued by linguistic anthropologist Webb Keane (2007) to constitute part of a more general Western modern perspective on the nature of language: the notion that the ‘signifying form’ of speech—the physicality of its sounds and gestures—is merely ‘a superficial garb laid upon more fundamental, if immaterial, meanings’ (64). While Keane’s immediate interest is the way this introduced ideology played out in the encounter between Calvinist missionaries and religious traditionalists in the former Dutch colony of Sumba, he notes that an analogous devaluation of form undergirds the assumptions about the nature of the sign that informed Ferdinand de Saussure’s model of the linguistic system. These assumptions include:

the radical distinction between signs and the world; and the doctrine of arbitrariness, which held that there are in principle no relations between signs and the world signified except for those established by the conventions of the system itself, commonly understood as a code shared among its users. As a result, what Saussure called *semiology* concerned only virtual types, never concrete tokens. Although language, for Saussure, did have form, which is the basis of the system of differences, it did not have materiality in any consequential respect. [Keane 2007:22]

As our mundane example from *The Name Dictionary* illustrates, and as Keane argues at length, the careful separation of speech form and meaning into their proper spheres is not only an analytical project, but also in some ways a moral one. ‘To attend too much to the form of the sign’, as opposed to its meaning, ‘is not simply an error; it is an expression of human arrogance’ (Keane 2007:66). Just as Protestant Christians on Sumba disapproved of traditional religious expression through formulaic speech rather than through spontaneous prayer because they felt only the latter was capable of conveying a speaker’s interior belief, appealing to sound as the basis for name assignment is held by the author of *The Name Dictionary* to be materialistic, meaningless, and inauthentic. In both cases we find a semiotic process is dismissed as insincere because it derives its rationale not from matters of meaning but from features of its form.

I discuss these cases of problematic motivation and mapping here although they come from outside the grammatical literature because they are central to the subject of this book: the logic of noun classification in the Arapesh languages of Papua New Guinea. Arapesh noun class assignment and agreement both rely to a remarkable extent on a principle that is all but identical to the lowly assonance method for pairing Jewish names: lexical items are categorized on the basis of their word-final sounds, and syntactic agreement is marked by copying those same sounds. Moreover, virtually every possible noun-final sound is represented in the paradigm of noun class and agreement markers, reflecting an interpenetration of phonology and grammar that some would reject as unacceptably unconstrained. My aim is to describe these sound-based patterns so as to reveal their naturalness and elegance, establishing their place in a typology of noun classification systems and drawing out their significance for any theory of grammatical architecture. But given the well-documented power of semantics to serve as a deeply embedded structuring principle for noun classification systems cross-linguistically, combined with the widely held assumption that phonological form must be morphosyntactically inert, there are points in the course of the exposition where it becomes necessary not just to describe how Arapesh noun classification works, but to defend in principle its ability to operate as concretely as it does. In short, our subject matter brings us head to head with the semiotic ideology Keane describes, and which Johanna Nichols (1986) has labeled the ‘Saussurean dogma’: the assumption that linguistic form belongs to a secondary, subordinate stratum of the semiotic order, the purpose of which is to code meaning, not to further motivate and organize linguistic form. If there is any grammatical phenomenon that provides grounds to question the Saussurean dogma, it is the kind of classification and agreement phenomenon described here.



Apart from some minor corrections, editorial improvements, and updates in the references, this work was completed in 1999 for my University of Chicago dissertation. Because my approach has been to analyze Arapesh noun classification in a theoretically neutral way, my arguments have remained essentially stable despite the passage of time. Regrettably, the same cannot be said for the Arapesh languages themselves, which have continued to decline in use as their speaker base shifts to Tok Pisin and English. Many of the Arapesh people with whom I consulted most closely during fieldwork have since passed away, taking their knowledge of their traditional language with them. I mourn the loss of these friends and continue to value what they so generously shared with me.

During my fifteen-month research stay in Papua New Guinea (December 1997 to March 1999) the people of Wautogik village, East Sepik Province adopted me as a *Wautogepierik*^w—as their sister, auntie, daughter, and granddaughter. I am indebted to them for their generosity and care. Especially deserving of thanks for their assistance with my research are Antonia Guaigu, †Joe Guaigu, †Clemen Hayin, †Lusia Hayin, Lausu Kwainger, †Pita Kwainger, Giamur Mujuwe, †Bernard Narokobi, Dorote Rahiria, †Matthew Wahiria, Antonia Sengu, Jacob Suonin, Scola Suonin, and †Arnold Watiem. Matthew Rahiria in particular spent hundreds of hours with me, despite debilitating illness, and is responsible for the vast majority of what I managed to understand about his extraordinary language. I must also acknowledge the special souls who enabled me reach Papua New Guinea, thrive while there, and maintain my ties to the community across the distance and subsequent years: Nick Araho, Bob Conrad, JoAnn Conrad, †Laura Martin, Andrew Moutu, †Bernard Narokobi, †Otto Nekitel, and Tony Nindim. I thank Emmanuel Narokobi and Vergil Narokobi for their interest in my work, and I applaud their commitment to stewarding the legacy of Wautogik village and the Abahinem clan. Bernard Narokobi's creativity, leadership, and insistence on the timeless value of Melanesian cultural ways were as much an inspiration to me as they were to the others who knew him. His memory is a blessing.

Material in Chapter 6 originally appeared in the article 'Theoretical Consequences of Literal Alliterative Concord', pp. 127-142, *CLS 31: Papers from the 31st Regional Meetings of the Chicago Linguistic Society. Volume 1: The Main Session*, copyright 1995 by the Chicago Linguistic Society; it is included here with permission. Material in Chapter 5 originally appeared in 'The Morphosyntactic Reality of Phonological Form', pp. 59-81 of the *Yearbook of Morphology 1997*, copyright 1998 by Kluwer Aca-

ademic Publishers. It is republished here with kind permission from Springer Science and Business Media B.V. The diagram in example (1) of Chapter 4 was copied with permission from George Lakoff's book *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, copyright 1987 by the University of Chicago.

Fieldwork on Arapesh was supported by a Fulbright-Hays Training Grant for Doctoral Dissertation Research Abroad (Award No. P022A70043), a Wenner-Gren Predoctoral Grant (Award No. 6156), and an NSF Dissertation Improvement Grant (Award No. SBR-9707681). I am grateful to these institutions for making the research possible, and to Stephen Leavitt and †Donald Tuzin for generously sharing their Arapesh fieldnotes with me. I thank Mark Aronoff, John Goldsmith, Gene Gragg, Richard Janda, and Jerry Sadock for their confidence and critical feedback during my years as a graduate student at the University of Chicago, when I was first finding my way into this project. My colleagues in linguistic anthropology at the University of Virginia, Ellen Contini-Morava, Eve Danziger, and Dan Lefkowitz, have been professional role models as well as mentors and dear friends. My debt to each one of them is immeasurable. I thank Farrell Ackerman, Bill Foley, and Andrew Spencer for their constructive criticism, and Emma Pease at CSLI for her help preparing the manuscript. Kelly Johnston's skill and patience are evident in the beautiful map. I am grateful to Amanda Sweeny for her keen eye and editorial judgment.

Nothing I have done over the past fifteen years would have happened—or happened so well—without the love and extraordinary assistance of my husband and colleague Ira Bashkow. This book is no exception. From advising me on writing research proposals to negotiating on my behalf with bus drivers in Papua New Guinea; from first introducing me to the art of fieldwork to staying home with the children so that I too could work, Ira has been both my closest partner and most reliable guide. I am humbled by the richness of his contribution to every aspect of my work and life.

Symbols and Abbreviations

Except for minor normalizations, the transcriptions used in this book remain faithful to the original sources. This has the advantage of facilitating the identification of examples and helps avoid projecting assumptions onto other authors' representations. However, the source I cite most heavily, Reo Fortune's 1942 grammar of Rohwim Arapesh, happens to be highly idiosyncratic in its transcriptions. The following is my best interpretation of Fortune's most frequently occurring symbols that depart from convention or that differ from those I adopt elsewhere. Superscript symbols here are printed in Fortune's grammar with under-ring. Further explanation is provided at various points in the text.

<u>Fortune</u>	<u>Interpretation</u>
<i>j</i>	palatal glide
<i>dz</i>	palato-alveolar affricate
<i>ø</i>	mid central vowel (often rounded)
<i>y</i>	high central vowel (often rounded)
<i>^h</i> or <i>^{uh}</i>	voiceless glottal fricative
<i>^u</i>	labialized release
<i>ⁱ</i> or <i>^y</i>	high vowel colored release

Fortune's glossing does not recognize the mood vowel that appears after the subject marker in Arapesh verbs; nor does it distinguish vowels that have a morphological function from those that arise in order to break up consonant clusters. His sub-word units therefore have extra material in them, such as the *a* and *o* in *fa-it-og^u*, where *a* is a realis mood marker and *o* is epenthetic. I follow Fortune in this non-standard practice here, since the

information is never critical to the argument and it helps keep the glosses uncluttered.

Forms taken from my own field notes are presented in roughly phonemic transcription. In Cemaun Arapesh, the symbols *c* and *j* represent voiceless and voiced palato-alveolar affricates, respectively.

The following abbreviations are used in the glosses:

1, 2, 3	first, second, third person; noun classes 1, 2, 3, etc.
I, II, III	Arapesh noun classes I, II, III, etc.
ABL	ablative
ADJ	adjective
ACC	accusative
AGR	agreement
AUG	augmentative
CNCT	connective
DAT	dative
DEF	definite
DIM	diminutive
F	feminine
GEN	genitive
IMP	imperative
INDEF	indefinite
IVΔ	intervocalic change
LOC	locative
M	masculine
N	neuter
NEG	negative
NOM	nominative
NP	noun phrase
OBL	oblique
PRF	perfective
PL	plural
POSS	possessive
PST	past
REL	relative
SG	singular
VOC	vocative

Introduction

1.0 Overview

This book offers a detailed analysis of the noun classification systems of the Arapesh languages, a relatively homogeneous family of languages spoken in northern Papua New Guinea. The Arapesh languages exhibit a structural feature which puts them at a typological extreme: an extensive set of morphological rules assigning nouns to classes not on the basis of their meaning, but rather on the basis of their concrete phonological form. Put simply, Arapesh categorizes nouns according to the way they sound. Arapesh has many class-assignment rules of the form $/Xy/ \rightarrow \text{class } Y$, where y represents the final phonological element of a noun, generally a single consonant, and the number of classes ranges in the teens. For example, Rohwim Arapesh nouns ending in the sound t regularly fall into a single noun class, so that the words *alit* ‘shelf’, *nokwat* ‘mouth’, *nybat* ‘dog’, and *unut* ‘bundle’ all take the same set of anaphoric and cross-referential agreement forms; that is, they require the identical adjective suffix, verb prefix, pronominal markers, and so on.

The existence of morphological systems utilizing form-sensitive rules like the Arapesh class-assignment rules bears on several central issues in the analysis of word structure. Perhaps most generally, it presents a clear example of a reflexive grammatical phenomenon, one in which grammatical form refers to other aspects of form, rather than to semantic information which it could be said to code. As such, evidence from Arapesh adds to the growing argument against what Nichols 1986, 1988 has called the ‘Saussurean dogma’, the widely held assumption that distinctions in linguistic form exist in order to encode meaning. A careful study of Arapesh noun class morphology also raises issues regarding the status of defaults, the determination of

morphemic status, the accessibility of phonological information to syntax, redundancy in lexical representations, and the relation between regular and irregular morphology.

In the remainder of this chapter, I discuss the problems inherent in any attempt to extend referential semantics or natural gender to cover all nouns in a gender or class system. Even the earliest philosophers who were concerned with the nature of language recognized class assignment rules referring to form, and modern studies of gender in French and German have demonstrated the formal basis of classification in more familiar languages. As we will see, what makes Arapesh typologically peculiar is the fact that its class-assignment rules refer to segmental phonology and that they do so in a remarkably transparent and exhaustive way: nearly every phonotactically possible noun-final consonant defines a unique class. Although most Arapesh noun class systems are large, it is not their size as such that makes them interesting. As Maho (1999:54) points out in a comparative study of Bantu languages,

The number of existing noun classes and/or pairings is not... a very good feature for typologizing noun class systems, at least not when looking for fine-grained categorizations. If a language has, say, 15 noun classes and at some point in its history reduces this number by one or two, for whatever reason, the noun class system as a whole is not likely to change in its basic grammatical functions. Some nouns are likely to be reallocated into new classes but we would still have a noun class system that works basically the same as it did before the change took place.

What is more important is the logic of the system. Only one Arapesh language, Weri, has changed in a way that shows phonological assignment to be fully disrupted.

In Chapter 2, I analyze the noun class system of the Rohwim Arapesh variety described by Fortune 1942 and draw consequences for morphological architecture. I will show that the model proposed by Aronoff 1994, which relies crucially on a set of hierarchically organized levels, is overly restrictive. The Arapesh patterns can be described much more naturally by treating noun class and plural realization not as structural levels within a morphological component of grammar, but rather as categories of information associated with nominal lexemes.

In Chapter 3, the Arapesh facts are brought to bear on the issue of how much information the lexical representations of word form must contain. Because many Arapesh nouns require partially predictable information to be listed in their lexical entries, the analysis in Chapter 2 makes a strong case for redundantly specified lexical entries along the lines of Bochner 1993, bringing fresh evidence to bear on a matter that is often decided by aesthetics and tradition.

In Chapter 4, noun class categories are investigated in relation to other morphological categories that have been discussed at greater length in the linguistic literature. The Arapesh form-based classes resemble semantically based classifier systems, more circumscribed cases of irregular morphology such as the English ablaut pattern (e.g. *sing* ~ *sang* ~ *sung*), and purely formal rule constellations in their internal organization, suggesting that common cognitive principles underlie these various forms of linguistic categorization. Chapter 4 also introduces the issue of how form-based and semantic class assignment rules are prioritized with respect to one another, a theme that runs throughout the next two chapters.

Chapter 5 considers one type of morphological category in detail, the default category. Morphological systems often exhibit a last resort rule that serves to incorporate forms falling outside the more typical generalizations. While to a great extent rule-like in its behavior, the Arapesh default plural *-ehas* also exhibits type-frequency effects of the kind associated with lexical networks. This suggests that regular and irregular morphology is distinguished by degree, and not by qualitatively distinct processing mechanisms, with one using symbolic rules to concatenate a variable with an affix and another consisting of overlap in a representational network.

Despite the existence of a native morphological default category, several varieties of Arapesh have innovated a new alliterative agreement pattern specifically for borrowed nouns that end in the sound *s*. This observation forms the basis for distinguishing between two types of formal class-assignment rules, those that refer to a noun's form by way of a morphological specification and those that refer to phonological form directly.

The topic of Chapter 6 is the possible interaction between the central grammatical components phonology, morphology, and syntax. The Arapesh agreement facts, bolstered by evidence from other languages, challenge even the weakest version of the lexicalist hypothesis, according to which phonology and syntax are prohibited from interacting directly. The Arapesh evidence suggests that syntactic rules are not universally 'phonology-free' (Zwicky and Pullum 1986a,b; Zwicky 1987; Miller, Pullum, and Zwicky 1997, Corbett 2010). That is, syntactic decisions cannot always be made without reference to phonological information.

Chapter 7 offers a brief conclusion.

1.1 The Arapesh Languages

1.1.1 Historical and Typological Orientation

The Papuan languages are the non-Austronesian languages indigenous to mainland New Guinea and its neighboring islands (see Wurm 1982; Foley

1986, 2000). The Arapesh languages form part of this geographically defined set. Arapesh belongs the Torricelli phylum (Laycock 1973, 1975). Torricelli characteristics include SVO word order, nominative-accusative syntactic alignment, relatively simple verbal morphologies, complex nominal systems involving a high degree of irregularity in patterns of plural formation, and, to varying degrees, phonologically-based noun classification and agreement (Foley 1986, 2000; see also Donohue 2006). The earliest typology of New Guinea languages, which placed heavy emphasis on morphological syndromes, classified Arapesh as ‘object-’ as opposed to ‘event-dominating’ because of the relative importance of nouns in its grammatical organization (Capell 1969). While the shared presence of noun classes does not necessarily imply that a group of languages is related (Terrill 2002), in Arapesh the status of the family is not in question. Both the Torricelli phylum and the Arapesh languages’ affiliation to it are supported by evidence from reconstructed pronominal forms (Ross 2005, n.d.).

The Arapesh family includes at least three major language groups: Mountain Arapesh or Bukiyip (2003 pop. 16,200), Southern Arapesh or Mufian (1998 pop. 11,000 all dialects), and Bumbita Arapesh or Weri (2003 pop. 4,340; see Lewis 2009). Nekitel 1985, 1986 carves out from this classification a fourth language, Abu’, a proposal justified by further lexicostatistical analysis (Gray, Retsema, and Hiley 2008). Comparative evidence suggests that Mufian, Weri, and Abu’ form a sub-grouping of Arapesh (Dobrin 2011). At the same time, given the considerable chaining of phonological, lexical, and grammatical features as one moves from village to village—a common distributional pattern in New Guinea—Arapesh is perhaps best understood as one long dialect chain on which this language classification is superimposed (Conrad 1978).

The names ‘Mountain Arapesh’ and ‘Southern Arapesh’ originate with Laycock 1973; Fortune 1942 uses only the term ‘Arapesh’ (this is the generic term for ‘people’ or ‘friends’ in the language). When possible I label language varieties after the local groups who speak them or by the village where they are spoken; in particular, I use ‘Rohwim’ to refer to the variety of Mountain Arapesh described by Fortune. In Dobrin 2001 I used a new label, ‘Arapeshan’, to refer to the language family as a whole, but despite ambiguity ‘Arapesh’ is well established in the linguistic and anthropological literature, so it is retained here. The glottonyms used by Arapesh-speaking people to refer to their own languages are frequently variants of the form *buk* or *buki*: *Bukiyip*, *Bukip*, *Buki*, *Abu’*; Bumbita Arapesh speakers refer to their language as *Weri* (as in many communities of the Torricellis, the language is referred to by its word for ‘no’).

1.1.2 Arapesh Varieties and Sources

This study is based on documentary sources for the following Arapesh varieties:

Coastal varieties of Mountain Arapesh

Fortune 1939, 1942 describes the Rohwim dialect spoken in present-day Woginara village. Gerstner's data (1939, 1963) represents the dialect spoken around a mission station then located near the village of But. My own fieldnotes represent the Cemaun dialect spoken in Wautogik village, on the northeastern-most border of the Arapesh area (see Dobrin 2001).

Bukiyip

Bukiyip is an inland dialect of Mountain Arapesh spoken on the southern side of the Prince Alexander Mountain range (Conrad 1978, 1987; Conrad and Wogiga 1991).

Abu'

Nekitel 1985, 1986 describes the variety of Abu' spoken in the village of Womsis. My fieldnotes document Hwamsək Abu'.

Mufian or Muhiang

These are the varieties of Southern Arapesh described by Alungum, Conrad, and Lukas 1978, and Matthews n.d.

Weri or Bumbita Arapesh

Weri is documented by Leavitt n.d.:a,b,c, as well as in my fieldnotes.

Data on several Arapesh varieties was collected by the author during 15 months of field research (1997-99) in East Sepik Province, Papua New Guinea. Material in this collection represents the coastal dialects of Mountain Arapesh spoken in Wautogik, Dogur, Woginara, and Balam villages; the Weru language; and the dialect of Abu' spoken in Hwamsək village (also spelled Hwamsuk; see www.arapesh.org).

Knowledge of the geographical relations among dialects is important for elucidating some of their linguistic differences. A map of the Arapesh language area noting key locations is presented in Figure 1 on page 8. Knowledge of the dates during which data on the various dialects was collected is likewise important, since the sociolinguistic scene in Papua New Guinea has changed significantly over the course of the past century, and the Arapesh varieties can only be understood and compared with these changes in mind.

The Rohwim dialect of Mountain Arapesh was documented by Reo Fortune on two research trips he made to the Arapesh region in the 1930s. Rohwim is spoken in the Prince Alexander Mountains just above the strip

of inland plain between Dagua and Matapau on the coast. Fortune distinguishes the dialect from its neighbors mainly on the basis of its consonantal inventory, and describes it as having thirteen classes of syntactic relevance, around half of them subsuming multiple plural subclasses. Fortune's description emphasizes the language's nominal morphology, but it also includes nearly 150 pages of glossed and abstracted Arapesh legends. Additional texts of traditional Arapesh oratory are included in Fortune's historical study of Arapesh warfare (Fortune 1939). I collected further vocabulary, grammar, and texts at Woginara, as well as at the nearby coastal village Dogur, where a closely related dialect is spoken. When the term 'Arapesh' is used here to refer to an individual language variety without further specification, it always indicates data taken from Fortune.¹

Another dialect of coastal Arapesh was described by the German missionary A. Gerstner, who conducted research in the But district between 1927 and 1937 (Burgmann 1963). There is no indication that Fortune and Gerstner were aware of one another's work. Gerstner describes seven more noun classes than does Fortune, but this is because he counts agreement with first- and second-person forms as a type of noun class agreement, whereas Fortune more accurately treats class as a property of the third-person only. Other differences between Fortune's description and Gerstner's apparently stem from the researchers' differential linguistic sensitivities: Gerstner was a native speaker of German, and this seems to have affected his perception of certain phonological distinctions, such as voicing in word-final position. Because this plays a crucial role in the selection of both agreement classes and plural forms, Gerstner's descriptions fail to identify significant regularities in the Arapesh system. Nevertheless, his documentation is useful in that it provides a way to verify some of the phenomena presented in Fortune's grammar. The vocabulary, grammar, and texts I collected at Balam village are readily recognizable as this or a closely related dialect.²

¹While it has some serious shortcomings (such as failing to indicate mood, a criterial grammatical category of the verb), Reo Fortune's work on Arapesh brilliantly captures the essence of the language's noun classification system. My appreciation for Fortune's Arapesh grammar and texts has led me to also revisit his work on Arapesh culture, which turns out to be just as insightful (see Bashkow and Dobrin 2007; Dobrin and Bashkow 2006, 2010a,b).

²Another early German missionary publication, Klaffl and Vorman 1905, provides a short (120 item) word list for a language they call *Kavu* (unfortunately the elicitations were cut short by the unexpected arrival of Klaffl's ship home). It was apparently collected in a coastal village at the western end of the Arapesh region, near the village of Matapau. *Kavu* seems to be a variant of the glottonym *Buki*, and the language is very similar to the But variety described by Gerstner. The description replicates certain features of Gerstner's grammar, including the suspicious neutralization of word-final obstruent voicing.

My own fieldnotes provide information on Cemaun, a previously undocumented variety of Mountain Arapesh. Cemaun is spoken in Wautogik and Kotai villages, at the northeastern border of the Arapesh dialect continuum in East Sepik Province, where Arapesh contacts Boikin, a language of the genealogically distinct Ndu family. This dialect is named after the traditional political alliance to which Wautogik belonged before warfare was outlawed by the German colonial administration in 1914.

Bukiyip is spoken between the towns of Yangoru and Maprik on the southern side of the Alexander Mountain Range. It has been studied intensively by Robert Conrad and his colleagues under the auspices of SIL International, with the practical aim of producing vernacular translations of Christian religious materials. The Bukiyip data was collected while Conrad was based in Bubuamo between 1971 and 1976. Conrad and Wogiga 1991 report an 86% stem-cognate count between Bukiyip and the coastal dialect of Arapesh. They list eighteen noun classes of syntactic relevance for Bukiyip. Additional data is included in Conrad's 1987 study of Bukiyip narrative discourse and a 1996 unpublished manuscript on Bukiyip noun classes. Occasional reference will also be made to data from a set of seventy-five unpublished Bukiyip texts (Conrad n.d.).

The dialects of Mufian are, from north to south, Supari, Balif, and Ilahita (Glasgow and Loving 1964; Alungum, Conrad, and Lukas 1978). Analyses here are based primarily on data from the central Balif variety spoken in the region around Albinama, and on Ilahita word lists, grammatical notes, and pidgin-glossed 'native stories' which were compiled by Shirley Matthews of the South Seas Evangelical Mission sometime prior to 1957. Also consulted were miscellaneous words, phrases, and Bible passages with Tok Pisin glosses recorded by anthropologist Donald Tuzin in fieldwork conducted in Ilahita around 1970. Alungum, Conrad, and Lukas 1978 identify seventeen noun classes in Mufian ('Mufian' is the spelling currently preferred by speakers).

Otto Nekitel, a linguist and native speaker of Abu' who hailed from Womsis village, was the first to identify Abu' as a language distinct from both Mufian and Mountain Arapesh. In his 1986 sketch of the Womsis Abu' nominal system, Nekitel described Bukiyip as 'fairly difficult' for him to comprehend. Nekitel describes nineteen noun classes in Abu'. Additional sources on Abu' noun classification include Nekitel 1984, 1985, 1992, 1998b as well as my own fieldnotes on the distinct Hwamsək dialect spoken 15-20 kilometers southwest of Womsis.

While the southwestern-most Arapesh language, Weri, clearly forms part of the Arapesh dialect chain, it also stands out against the backdrop of Arapesh languages in that while it maintains cognate singular and plural forms, it assigns nouns to agreement classes on purely semantic grounds.

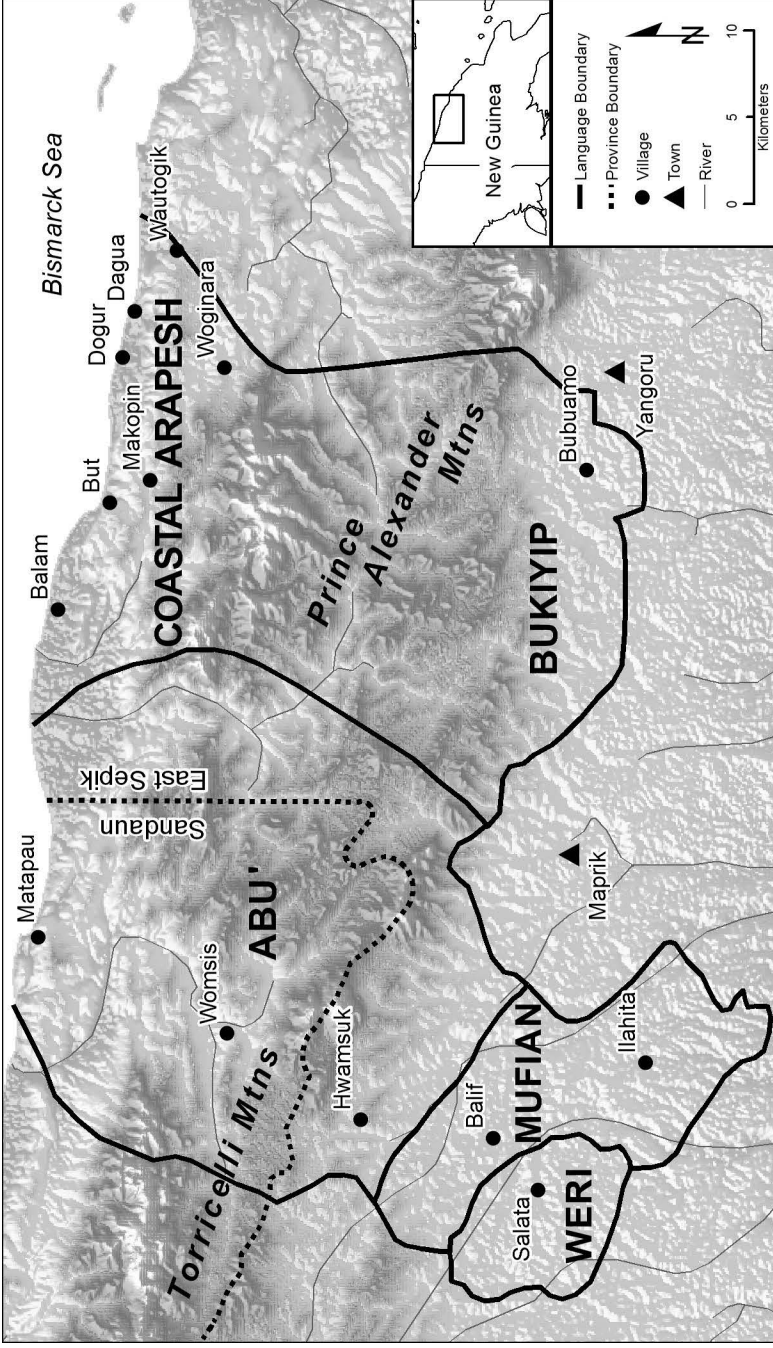


Figure 1. The Arapesh Language Area

The Weri materials discussed here add to the information collected by anthropologist Stephen Leavitt (n.d.:a,b,c).

As with so many other of the world's minor languages, several of the Arapesh dialects are severely endangered, with little chance of surviving beyond the current generation of speakers. The advancement of Western products, lifestyles, and values is unrelenting, and the spread of Tok Pisin, the primary lingua franca in Papua New Guinea and the rural prestige language, has already reached the point where all but the oldest speakers of Arapesh are not only fluent in Tok Pisin, but moreover grew up Tok Pisin-Arapesh bilingual. Unfortunately, this situation is increasingly common throughout the Sepik region (Foley 1991, Kulick 1992, Aikhenvald 2004). Nidue 1990 described the Mountain Arapesh village of Makopin as triglossic in Arapesh, Tok Pisin, and English, with an increasingly circumscribed role for the vernacular as inter-group contact increases. Nidue projected that Arapesh would die in Makopin within fifty years. Nekitel 1984 wrote that Abu' was ceasing to be transmitted to the current generation of language learners in his native village Womsis. At the time of my field research in the late 1990s, few coastal Arapesh villagers under the age of about thirty were fluent speakers; what knowledge they had of their vernacular was primarily passive. When spoken to in Arapesh, they would routinely respond in Tok Pisin, which is now a universal native tongue throughout the region. School-age children today will rarely produce the standard greetings 'good morning' or 'good night' in their ancestral language, either because they do not know them or because their lack of confidence with the vernacular makes them hesitant to use them. A recent SIL survey in the western Arapesh territory confirms that these trends continue (Gray, Retsema, and Hiley 2008).

1.2 Noun Categorization and Word Form

The terms 'noun class' and 'gender' are used here interchangeably to refer to noun categorizations that are reflected in the agreement patterns of associated words.³ Reference to grammatical properties outside a noun itself is a widely agreed upon aspect of this definition, and such an agreement criterion has been assumed in the foundational studies of gender and noun classification (e.g. Fodor 1959, Greenberg 1978). Such a definition rules out the possibility of calling, for example, the set of English abstract nouns suffixed with *-tion* a gender, since this property is not reflected in agreement

³Not all discussions of noun classification follow this usage; some distinguish natural gender (masculine, feminine, neuter) from non-sex-based classes (see, e.g., Beard 1995:103-111, Aikhenvald 2000, Mel'čuk 2006).

(Corbett 1991). But there is less consensus on what exactly should count as agreement. Some authors (e.g. Dixon 1982) would exclude pronominal agreement or anaphora from the definition, which would thus remove the English pronouns *he*, *she*, and *it*, and with them the entire English language, from the arena of noun categorization. Following Corbett I will assume here that anaphoric agreement falls within the purview of noun class agreement.

The grammaticization of sex, animacy, and/or humanness in many languages is not in itself surprising, given the importance of these features of human experience.⁴ But the principles by which classification is extended to words for inanimate objects have concerned students of language since even before the time of Aristotle. The non-natural quality of noun classification can be seen especially clearly in the conflicting gender assignments of nouns referring to the same object in different languages. In Spanish, for example, a table is feminine (*la mesa*), as is the moon (*la luna*), while in German the nouns referring to these objects are masculine (*der Tisch* and *der Mond*, respectively). Such differences in assignment may even characterize words that are cognate in two languages. So, for example, the noun ‘tooth’ is feminine in French (*la dent*) but masculine in Spanish (*el diente*). Why are tables and the moon feminine in Spanish, but masculine in German? Why are teeth masculine in Spanish, but feminine in French? Leonard Bloomfield’s famous answer to the question of why genders differ among languages stands as the null hypothesis: for inanimate referents, ‘the gender categories of most Indo-European languages... do not agree with anything in the practical world.... There seems to be no practical criterion by which the gender of a noun in German, French, or Latin could be determined’ (1933: 271, 280).

Sometimes cultural knowledge can supply a crucial link. In the Australian aboriginal language Dyirbal, for example, the word for ‘moon’ falls into the masculine category due to its anthropomorphic status as a male: in Dyirbal mythology the moon is the sun’s husband (Dixon 1972, 1982; Lakoff 1986:92-104). But such associations are not always forthcoming. Consciously, at least, tables are no more female to speakers of Spanish than they are male to speakers of German, though there is evidence that language-particular gender categorizations can affect non-verbal cognition independently of culture (Boroditsky, Schmidt, and Phillips 2003).

A partial correlation between male vs. female sex and masculine vs. feminine grammatical form was understood as early as the fifth century BCE, when Protagoras associated the genders of Greek words with the natures of the objects they stood for (see Robins 1951, 1990). Exceptions to

⁴The fundamental grammatical distinction is that ‘between persons, that is, essentially human beings perceived as agents, and the rest of the universe’ (Dahl 2000:100).

semantic generalizations were recognized by the ancients as well; Protagoras even went so far as to suggest that the gender of the Greek nouns *menis* ‘anger’ and *pelex* ‘helmet’ be changed from feminine to masculine to better accord with these terms’ association with males rather than females!

But even as they sought to explain gender in terms of meaning, the earliest Western grammarians noted a correlation between gender and nominal form. In his *Poetics* (at 1458a) Aristotle lists noun-final sounds which are typically associated with gender in Greek: ‘All [nouns] ending in N [n], P [r], Σ [s], or in the two compounds of this last Ψ [ps] and Ξ [ks], are masculines. All ending in the invariably long vowels, Η [e:] and Ω [o:], and in Α [a] among the vowels that may be long, are feminines.’

French and German, two of the languages cited by Bloomfield to illustrate the arbitrariness of gender, have been shown to have highly systematic gender systems in which aspects of word form are significant predictors. In French, there are semantic rules that assign nouns denoting humans to masculine and feminine genders on the basis of sex (e.g. *père* M ‘father’, *mère* F ‘mother’). For some lexical fields, there are also covert semantic classifications akin to those used in classifier systems (e.g. upright and flexible attributes predict masculine whereas horizontal and rigid predict feminine; see à Beckett 2010). But there are also morphological predictors for complex nouns (e.g., deverbal nouns formed with the suffix *-ation* are feminine), as well as rules referring to the word-final phonological element, such as masculine gender assignment to nouns ending in *η*, *τς*, and *δζ* (Tucker, Lambert, and Rigault 1977). Although many of the phonological assignment rules admit exceptions, they nevertheless form an important component of the native speaker’s linguistic competence, predicting the categorization of both native lexical items and nonce words.

Gender in German has similarly been shown to be partially predictable from a noun’s morphological and phonological form (Zubin and Köpcke 1981, Köpcke 1982). For example, feminine gender correlates exceptionlessly with certain derivational suffixes (e.g. *Schönheit* F ‘beauty’, *Freundschaft* F ‘friendship’, *Schöpfung* F ‘creation’); and it tends to be assigned to nouns that select the *ən* plural (e.g. *Perle* F ‘pearl’ ~ *Perlen* ‘pearls’). Nouns containing long vowels are very rarely feminine (e.g. *Gruß* M ‘greeting’, *Brot* N ‘bread’). Nouns with initial or final consonant clusters have a significant tendency to be masculine, as do monosyllabic nouns.

From a psycholinguistic perspective in which language is recognized as partly perceptual, and not an exclusively cognitive system, information about a word’s phonological form has been shown to play an important role in determining high level lexical categories in addition to lexical sub-classes like genders. Kelly 1992 reviews literature establishing that major lexical classes of English (noun vs. verb, open vs. closed) correlate with a range of

phonological variables, including the location of word stress, syllable count, phonetic duration, and vowel quality. These correlates, like the phonological correlates of gender mentioned for French and German above, are moreover relied upon as cues in language acquisition, production, and comprehension. As Cassidy, Kelly, and Sharoni 1999 demonstrate, speakers use these same phonological features to establish the gender associations of American English names.

The observation that class-assignment rules may be grounded in a noun's sound shape has serious implications for our understanding of the nature and functioning of language. Linguistic analysis continues to be driven by what Nichols 1986, 1988 has dubbed the 'Saussurean dogma': the assumption that 'content determines, or limits, or motivates, form, and not versa' (1986:142). Lass 1990:100 refers to this as the 'semiotic fallacy'. The tenacity of such an assumption is understandable: what else is linguistic form for, if not to encode meaning? But it seems that meaning can be somewhat more abstract than we usually take it to be. In an article entitled '(What) Do Swahili Noun Class Markers Mean?' Contini-Morava 2002 argues that although the semantic associations of most of the Swahili class markers are disparate, class marking nevertheless serves a useful signaling function in that it partitions the set of noun stems into subsets for the purpose of establishing discourse coreference. In that case, their meanings must be understood as 'processing aids rather than notional fractions of the message' (9). This is reminiscent of the conclusion drawn by Aronoff (1976:15) that 'what is essential about a morpheme... [is] not that it mean, but rather merely that we be able to recognize it.'

The case against the Saussurean dogma made by Nichols 1986 appeals to examples of formal linguistic codings that are 'reflexive'. In languages that distinguish verb inflection principally into active and stative subtypes, head-marked verbs of a certain semantic category, namely, bodily functions such as 'sneeze', are consistently assigned to the more marked or minority class of intransitives, regardless of what else the class contains. Thus, Nichols argues, the categorization of bodily-function verbs in head-marking languages is reflexive: it does not directly reflect their meaning, but is determined by an abstract formal property of the grammatical system in which they participate, the relative markedness of a formal subclass.

Other arguments for meaningless morphology come from morphological generalizations of a sort Janda 1982 and Janda and Joseph 1986, 1992, 1999 call 'meta-redundancy rules', 'meta-templates', or 'rule constellations'. Meta-rules fail to correspond in a systematic way to any morphosyntactic or morphosemantic category, and thus cannot be assigned content or attributed a signaling function. In other cases rule constellations arise not because of irregular semantic or syntactic properties, but because the rules

they subsume cannot be formally unified. The meta-rule for German umlaut provides a good example. There are multiple umlaut rules in German, each bringing about in part the same formal change. Yet different umlaut rules have somewhat different structural descriptions. For example, noun-pluralizing umlaut may apply to the diphthong *au* as in *Haus* ~ *Häuser* ‘house’ ~ ‘houses’, whereas the umlaut rule marking comparative adjectives may not apply to this sound; thus, *laut* ~ *lauter*/**läuter* ‘loud’ ~ ‘louder’. Umlaut rules are also differentiated by the fact that they apply in combination with distinct sorts of other morphological markers, in addition to applying on their own, e.g. *Ball* ~ *Bälle* ‘ball’ ~ ‘balls’, *Dach* ~ *Dächer* ‘roof’ ~ ‘roofs’, *Vater* ~ *Väter* ‘father’ ~ ‘fathers’. Furthermore, different umlaut rules have different morphosyntactic functions, marking, e.g., the superlative of adjectives *arm* ‘poor’ ~ *ärmst* ‘poorest’ as well as person on verbs *fall* ‘1.SG.fall’ ~ *fällst* ‘2.SG.fall’. Yet all these processes share a coherent formal core, the fronting of a back vowel. Rule constellations like umlaut are thus purely formal, partial generalizations that are not inherently associated with any category of meaning.

Like meta-rules, stems and inflectional classes—the arbitrary paradigms of declensions in nouns and conjugations in verbs—can also be seen as contributing to the overall organization of a morphological system even though their meanings can at best be defined in a formal way (Aronoff 1994, Carstairs-McCarthy 1994:741). A simple example is the English past participle. What signaling function does such a grammatical construct have? It can hardly be said to directly express a meaning; instead, it is an abstract category that mediates between particular verb stems (*walk*, *leave*, *give*, etc.) and their surface realization in a disjoint set of morphosyntactic constructions (the passive and perfect). Or consider the stem-final theme vowel of Latin verbal morphology. ‘In itself, it has no significance. It is empty. Nonetheless, it is not useless’ (Aronoff 1994:46). Reference to the theme vowel serves as a cue for assigning particular verb stems to the appropriate paradigm. This kind of system-dependent formal interpredictability can be extremely elaborate, a point Blevins 2006 illustrates with the weak and strong stems of Saami and Estonian noun declensions (see also Stump 2001). Aronoff 1994 calls purely formal generalizations of this kind ‘morphomic’. Although morphomic patterns do not convey meanings of their own, they nevertheless can be seen to play a critical role directing traffic in complex morphological systems.

The present work confronts the tension that exists in a family of morphological systems that exhibit the abstract, meaningless character of autonomous morphology and, almost paradoxically, a simultaneous dependence on another category of lexical information, phonological form. The Arapesh noun classes systematically categorize nouns according to

their word-final consonants, offering an especially clear case of a reflexive coding in which form organizes other aspects of form, and does not simply reflect meaning or realize content. Arapesh thus provides further justification for granting morphology a greater degree of autonomy from meaning than is traditional in morphological theory. At the same time, the Arapesh noun class systems make a compelling case for stating morphological generalizations directly in terms of segmental phonology. To resolve this tension between concreteness and abstractness in a satisfying way within a general descriptive framework is the goal of this book.

1.3 The Phonological Basis of Arapesh Noun Classification

1.3.1 How is Arapesh Special?

If reference to the form of a noun is common in class assignment as suggested above, then something further must be said in order to distinguish Arapesh from other languages that have noun class systems based partly in noun form, such as French and German. One consideration is the number of classes into which nouns in a system are divided. French has two genders and German has three, but the Arapesh varieties have upwards of thirteen syntactically significant noun classes. Yet sheer size is an insufficient basis for distinguishing Arapesh from more familiar languages with large noun class systems, such as those of the Niger-Congo family. Bantu languages in particular are often cited in the literature on noun classification as the prototype of large, form-based systems; Swahili, for example, is usually described as having seven or more noun classes.⁵ The Swahili classes encode animate, diminutive, and augmentative semantic features that can in some cases override the class assignment expected on the basis of form, but a Swahili noun's class is otherwise determined by its prefix. The word 'prefix' here is crucial: Swahili nouns contain an identifiable structural element to which classification refers—an affix that is morphologically distinct from the noun root. Arapesh nouns arguably do not. The Swahili noun classes are therefore associated with phonological form only indirectly due to the fact that nouns are structured in this way.

I should underscore what is being claimed here for Swahili. It is not the case that the class of a given Swahili noun is always marked overtly by a prefix. Among native nouns, there are some that are prefixless on the surface. For example, the prefixes of classes 5 and 9 have \emptyset allomorphs (1):

⁵Estimates like this are bound up with decisions about whether class is a property of lexemes or of individual inflected forms. For Swahili the former assumption is justified by the fact that classification typically has simultaneous consequences for the agreement behavior of both singular and non-singular noun forms (see Corbett 1991).

- (1) Class 5: *ji-cho* ‘eye’ vs. *elezo* ‘explanation’, *wingu* ‘cloud’
 Class 9: *n-dege* ‘bird’ vs. *taka* ‘dirt’, *shingo* ‘neck’

There are also many instances of nouns that have been borrowed into Swahili from other languages without undergoing any morphological adaptation at all. Indeed, this is the most common strategy for incorporating English vocabulary into the Swahili class system (Contini-Morava 1994).⁶ In these cases, as well, noun class is not associated with any identifiable prefix (2).

- (2) Class 5: *gita* ‘guitar’, *trekta* ‘tractor’, *blekibodi* ‘blackboard’
 Class 9: *foto* ‘photograph’, *fyuzi* ‘electrical fuse’, *jela* ‘jail’

Thus, the set of correlations between a Swahili noun’s initial segment(s) and its assignment to a particular agreement category—i.e. the set of form-based class-assignment rules—is much smaller than the set that would be possible on the basis of the noun-initial phonology alone. The major Swahili noun prefix patterns are listed in (3) (see Welmers 1973).

(3)	<u>Class</u>	<u>SG ~ PL Prefix</u>	<u>SG ~ PL Noun</u>	
	1 ~ 2	<i>m ~ wa</i>	<i>m-tu ~ wa-tu</i>	‘person’
	3 ~ 4	<i>m ~ mi</i>	<i>m-kate ~ mi-kate</i>	‘loaf’
	5 ~ 6	<i>ji ~ ma</i>	<i>ji-we ~ ma-we</i>	‘stone’
	7 ~ 8	<i>ki ~ vi</i>	<i>ki-su ~ vi-su</i>	‘knife’
	9 ~ 10	<i>N ~ N</i>	<i>n-goma ~ n-goma</i>	‘drum’
	11 ~ 10	<i>u ~ N</i>	<i>u-bao ~ m-bao</i>	‘plank’
	15	<i>ku</i>	<i>ku-imba</i>	‘singing’

Apart from the productive semantic predictors of noun class mentioned above, Swahili class-assignment is based on a noun’s pattern of prefixation in the singular and plural rather than on its meaning. It is for this reason that Corbett 1991:43-49 treats Swahili genders as belonging to the set of ‘formal gender systems’, and I will follow him in assuming that a Swahili noun’s

⁶The assignment of the nouns *kiplefti* (British English ‘roundabout’, from English ‘keep left’) and *kitabu* ‘book’ to the Swahili *ki-* class are frequently cited examples of phonological form influencing class assignment in borrowing. However, form-based assignment of loans in Swahili is the exception rather than the rule. As Whitely 1967:170 noted long ago:

The only words allocated to 7/8 are those with an initial /ki-/ element in the English e.g. *kiplefti* Roundabout; *kilabu*, Club; *kisi*, Kiss but even in these cases there is a tendency towards re-allocation to 9/10 e.g. *kisi moja*, *kilabu nyingi*.... If present tendencies are maintained and the number of loans continues to increase at the present rate, in a few years time a considerable proportion of the common vocabulary of Tanzanians will comprise nominals in two morphological classes only [5/6 and 9/10, the two classes with prefixless allomorphs].

class (as well as its plural) is generally predictable from its prefix in the singular. In that case there are only six formal predictors of gender in Swahili, each of them at most a single syllable long: *m-*, *ji-*, *ki-*, *N-*, *u-*, and *ku-*.⁷ There is no constraint of Swahili phonology that prevents a noun from beginning with many other possible syllables of the language, for example, *to*, *e*, *pe*, or *fa*, and there in fact exist native Swahili nouns beginning with precisely these sounds: *tofali* ‘brick’, *elezo* ‘explanation’, *pera* ‘guava’, and *shangazi* ‘aunt’. Yet these do not constitute noun prefixes, there is no class assignment rule in Swahili referring specifically to these initial sounds or sequences, and there is no tendency for Swahili speakers to form new agreement patterns for borrowings with novel initial sequences.

In Arapesh, by contrast, the shape of a given noun is determined essentially freely by language-specific principles of phonology, and the assignment of a noun to a particular class proceeds consequently on the basis of its phonological form. Thus, the large number of classes in Arapesh is intimately bound up with the fact that the class categories are rooted directly in nominal phonology. Most of the noun classes refer to the quality of the noun-final segment, typically a consonant.

It is more usual for noun classification systems to refer to phonological form at a level of description either higher or lower than that of the segment. In the East Cushitic language Afar, for example, there is a condition on assignment that specifies the category of ‘consonant-final’ nouns, thus referring to phonological form at a relatively high level of description that subsumes many more phonological subcategories. In Afar, consonant-final nouns are nearly always masculine, whereas vowel-final nouns may be either masculine or feminine, hence: *dumam* M ‘nose ring’ and *daas* M ‘booth’, vs. *dáyla* M ‘medicine’, or *cagáya* F ‘hot season’ (Parker and Hayward 1985). The Afar rule does not distinguish between nouns that are *r*-final, *m*-final, *s*-final, and so on, as the Arapesh gender-assignment rules do. In Arapesh, only vowel segments are referred to across qualities.

In other languages, phonologically sensitive gender rules refer to natural classes of sounds based on features below the level of individual segments. In the Kru language Godié, for example, there are three genders, each referring to a noun-final phonological element (see Marchese 1986, 1988 for the agreement facts). Godié nouns are always vowel-final, and what determines gender is which of three values along a single phonological dimension characterizes a noun’s final vowel: front (e.g., *gwe* ‘chimpanzee’), central (e.g., *kəmə* ‘crab’), or back (e.g., *zuzu* ‘spirit’). Arapesh class assignment differs

⁷The prefixes exhibit allomorphic variation depending on the form of the stem to which they attach.

from assignment in Godié in that it typically refers to the level of individual consonant phonemes, disregarding most featurally-defined natural classes.

1.3.2 Abu'

The Womsis variety of Abu' as described by Nekitel 1986 illustrates the phonologically exhaustive nature of noun class assignment in Arapesh. The Abu' phonemic inventory is summarized in (4) below. The sound *p* has entered the language only recently through Tok Pisin borrowings, and the apostrophe denotes a glottal stop. Nekitel 1986 assumes the inventory to contain fifteen native consonantal phonemes and five vowels. Nekitel 1998 lists an additional phoneme *x^w* for his dialect that is not mentioned in the phonology or orthography of the earlier publications; this accounts for his earlier division of *h*-final nouns into two classes, V and XIX.⁸ The primary Abu' classes posited by Nekitel 1986 are listed in (5).

(4)	<u>Vowels</u>		<u>Consonants</u>	
	<i>i</i>	<i>u</i>	(<i>p</i>)	<i>t</i> <i>k, k^w</i> ' ,
	<i>e</i>	<i>o</i>	<i>b</i>	<i>d</i>
		<i>a</i>	<i>f</i>	<i>s</i> <i>h (h^w)</i>
			<i>m</i>	<i>n</i>
				<i>l</i>
				<i>r</i>
			<i>w</i>	<i>y</i>

(5) Abu' Noun Classes

<u>Class</u>		<u>Noun</u>	<u>Adjective</u>	<u>Verb</u>	
I	SG	<i>aleman</i>	<i>-neri</i>	<i>n</i>	'man'
	PL	<i>alemam</i>	<i>-mi</i>	<i>m-</i>	
		<i>mohun</i>			'male in-law'
		<i>mohulihim</i>			
II	SG	<i>numata'</i>	<i>-'i</i>	<i>kw-</i>	'woman'
	PL	<i>numatawa</i>	<i>-weri</i>	<i>w-</i>	
		<i>a'u'</i>			'wife'
		<i>awuw</i>			

⁸In Hwamsək Abu' word-final *h^w* is phonemic and corresponds to a distinct noun class taking alliterative agreement with *h^w*. The examples Nekitel uses to justify classes V and XIX in his chart for Womsis Abu' belong to different classes in Hwamsək Abu': *manduh^w atuh* 'vine one' (final phoneme *h^w*); *iyah atih* 'path one' (final phoneme *h*).

(5) Abu' Noun Classes (cont.)

<u>Class</u>		<u>Noun</u>	<u>Adjective</u>	<u>Verb</u>	
III	SG	<i>baraka</i>	<i>-keri</i>	<i>k-</i>	'head'
	PL	<i>barakas</i>	<i>-kas</i>	<i>s-</i>	
		<i>lawak</i> <i>lawas</i> ⁹			'tree'
IV	SG	<i>aulaf</i>	<i>-fi</i>	<i>f-</i>	'house'
	PL	<i>aulas</i>	<i>-si</i>	<i>s-</i>	
		<i>bahiataf</i> <i>bahiatas</i>			'scaly river fish'
V/XIX	SG	<i>maduh</i>	<i>-hi</i>	<i>h-</i>	'vine', 'rope'
	PL	<i>madulih</i>	<i>-lihi</i>	<i>h-</i>	
		<i>iyah</i> <i>iyokwih</i>	<i>-hi</i>		'road', 'path'
VI	SG	<i>bakel</i>	<i>-kili</i>	<i>l-</i>	'stick'
	PL	<i>bakuh</i>	<i>-kuhi</i>	<i>h-</i>	
		<i>aul</i> <i>akuh</i>	<i>-li</i>		'mosquito', 'eel'
VII	SG	<i>utam</i>	<i>-mi</i>	<i>m-</i>	'stone'
	PL	<i>utaba</i>	<i>-beri</i>	<i>b-</i>	
		<i>adum</i> <i>aduba</i>			'right side'
VIII	SG	<i>abom</i>	<i>-mi</i>	<i>m-</i>	'breadfruit seed'
	PL	<i>abef</i>	<i>-fi</i>	<i>f-</i>	
		<i>subarum</i> <i>suberif</i>			'grass skirt'
IX	SG	<i>du'it</i>	<i>-teri</i>	<i>t-</i>	'mountain'
	PL	<i>disuk</i>	<i>-ki</i>	<i>k-</i>	
		<i>al'ut</i> <i>alsuk</i>			'dusk'

⁹This plural is extrapolated from my field notes on the Hwamsək dialect.

(5) Abu' Noun Classes (cont.)

<u>Class</u>		<u>Noun</u>	<u>Adjective</u>	<u>Verb</u>	
X	SG	<i>alhuab</i>	<i>-bi</i>	<i>b-</i>	'egg'
	PL	<i>alhuabis</i>	<i>-bisi</i>	<i>s-</i>	
		<i>uwab</i>			'night'
		<i>uwabis</i>			
XI	SG	<i>nubul</i>	<i>-bili</i>	<i>b-</i>	'stomach'
	PL	<i>nulub</i>	<i>-libi</i>	<i>b-</i>	
		<i>wabul</i>			'village'
		<i>walub</i>			
XVII	SG	<i>awata</i>	<i>-tari</i>	<i>t-</i>	'chicken'
	PL	<i>awataawk</i>	<i>-kwaki</i>	<i>kw-</i>	
		<i>bata</i>			'bamboo'
		<i>bataawk</i>			
XVIII	SG	<i>dubaun</i>	<i>-ni</i>	<i>n-</i>	'lobster'
	PL	<i>dubaub</i>	<i>-bi</i>	<i>b-</i>	
		<i>ehen</i>			'cold river fish'
		<i>eheb</i>			

Some Abu' nouns are classified by virtue of their semantics. As in the other varieties of Arapesh, there is one class that is exclusively masculine, Nekitel's class I, and another that is partially feminine, which he calls class II; the latter includes not only nouns referring to females persons, but many nouns with diverse other referents as well. Nouns in these semantic categories are also typically associated with a particular noun-final phonological form. For class I it is *n* or *na*. Most class II nouns end in a glottal stop.¹⁰

But the other nine Abu' classes do not have consistently associated semantics; for most Abu' nouns, class is determined on the basis of noun form. Closed syllables do not appear word-internally in the native Abu' vocabulary, but in word-final position syllables may be either open or closed. And for nearly every word-final consonant there is a corresponding form-based class assignment rule, i.e., a generalization like the following that refers specifically to that sound, where the arrow in the rule indicates a relationship of prediction.

¹⁰The numbering reflects the original source (Nekitel 1986). Gaps reflect classes posited for reasons not sanctioned here, or else classes for which Nekitel provides insufficient justification.

(6) $/XC_i/ \rightarrow \text{Class } C_i$

The argument that the Abu' classes are tied directly to the language's phonology runs as follows: for every member of the Abu' phonemic inventory that is phonotactically possible word-finally, there is a corresponding class assignment rule referring to that sound in noun-final position. For any Abu' phoneme that is not directly referenced by a class assignment rule, it should be possible to show that there is a crucially phonological constraint affecting the distribution of that sound, such that it is prohibited from appearing in the relevant class-defining position. As will be discussed in Chapter 2, the fact that these rules reference singular noun forms is not accidental.

This argument does not require that these be the only sorts of classification rules in Abu'. In addition to the two semantic assignment rules mentioned above, there are several rules referring to more than a single phonological segment. The assignment of class XI involves one such rule: class XI nouns end not just in a consonantal segment, but in the final CVC-sequence *bul*. There are at least three diminutive suffixes that take distinct agreement patterns and thus constitute minor classes in their own right: *-ikin* (as in *alimilikin* 'small bird'), *-tawas* (as in *lehitawas* 'small portion of sago'; cf. *lehin* 'sago'), and *-ikil* (as in *nikamikil* 'small taro'; cf. *nikam* 'taro'). Because they depend on the presence of suffixes that have a consistent meaning and can attach to an independently justified root, the diminutives have a clear basis in morphology, as opposed to phonology. Finally, a morphological default rule provides for agreement with nouns unaccounted for by any of the regular assignment principles (the singular default agreement marker is *n*). But morphologically defined classes such as these can coexist with a phonological class-assignment principle in the system. It is in fact one of the basic conclusions of this study that the existence of morphologically based class assignment provides the typological prerequisite for phonological class assignment. That is, as will be discussed in Chapter 5, there is a universal hierarchy such that the existence of phonologically based class assignment rules implies the existence of morphologically based rules in the language.

The first qualification to the claim that Abu' class-assignment refers directly to noun-final phonological form involves the role of final vowels. While vowels are phonologically admissible in word-final position (as in *niboa* 'two days ago', *emi* 'who', *ene* 'this.M.SG'), the only vowel which is found finally on native singular nouns is *a*. Nekitel specifically lists two classes making reference to this vowel (classes III and XVII), but it cannot be the noun-final vowel per se that is targeted; otherwise the two sets of nouns would be classified together. Comparative evidence and Abu' agreement

patterns suggest that what differentiates these two classes is not the noun-final vowel but the preceding consonants, *k* and *t*. Consider the following singular-plural pairs from cognate classes in Abu' and Rohwim Arapesh (7). There is some overlap in the vocabulary assigned to these classes in the two languages.

(7)	a.	Arapesh III	<i>aijag</i> ~ <i>aijas</i>	'leg'
			<i>berag</i> ~ <i>beragas</i>	'head'
		Abu' III	<i>baraka</i> ~ <i>barakas</i>	'head'
			<i>akiaka</i> ~ <i>akiakas</i>	'ant'
	b.	Arapesh XI	<i>kweit</i> ~ <i>kweitog^u</i>	'sago slicer'
			<i>abut</i> ~ <i>abutog^u</i>	'bamboo container'
		Abu' XVII	<i>bata</i> ~ <i>batawk</i>	'bamboo'
			<i>ahata</i> ~ <i>ahatawk</i>	'marsh'

Although Nekitel mentions two true *t*-final Abu' nouns, *du'it* 'mountain' and *al'ut* 'dusk', they have unexpected plurals (*disuk* and *alsuk*, respectively), and it is questionable whether these are systematic enough to warrant positing a noun class for them alone, as Nekitel does. Nekitel posits no category of *k*-final nouns, though in his text the word *lawak* 'tree' appears with a class III agreement marker, suggesting assignment to class III.¹¹ The special morphological significance of the consonants in *Ca#* sequences is also supported by the fact that the agreement markers for both *Ca* classes echo the *C*, as in all other classes. It seems, then, that the noun-final vowel *a* is not directly engaged in Abu' class marking. This pattern is replicated in all the southwestern Arapesh varieties.

Where there is no final consonant for a class assignment rule to refer to, as in the case of a vowel-final borrowing lacking any semantic features that would point to a particular classification, e.g. *redio* 'radio' or *ti* 'tea', a noun takes a default agreement marker *n* rather than a marker depending in some way on the noun's phonological form (*redio ene* 'radio this'). The following noun-final segments thus each play a direct role in assignment: *t*, *k*, *'*, *b*, *m*, *n*, *f*, *l*, *h*, and *h^w* (or *x^w*). The fact that the following members of the phonemic inventory do not participate in class assignment remains to be accounted for: *d*, *y*, *w*, *k^w*, *r*, and *s*.

¹¹In justifying his phonemic analysis Nekitel provides an example of another true *t*-final noun, *nubat* 'dog', which is a regular and frequent member of the cognate *t*-final class, class XI, in Rohwim Arapesh. This form also occurs in Hwamsək Abu', supporting the notion that word-final *ta* and *t* are not morphologically distinct in Abu'.

The consonants *d*, *y*, *w*, *k^w*, and *r* are excluded for a very simple reason: none of these sounds can appear noun-finally in Abu'. This is not due to a morphological condition that all nouns must be realized with one of a fixed set of class markers, but rather to independent phonological conditions on the shape of a possible Abu' word.

Nekitel states that no Abu' nouns end in *d*, though he provides no explanation for this gap. Taking the rest of Abu' phonology into account, such a restriction could be attributed to a constraint against voicing word-finally: before the recent introduction of voiceless *p* from Tok Pisin, *d* would have been the only distinctively voiced obstruent in the Abu' inventory, and so might be expected to show a defective distribution. With *d* ruled out of word-final position by a phonotactic constraint, the fact that this sound fails to participate in the class assignment system is not only unexceptional but predicted. Independent conditions on Abu' phonology prevent it from doing so.¹²

The absence of the glides *y* and *w* from word-final position is a function of the realization of vowel features in the syllable. Nekitel 1986:182 tentatively suggests that high glides are not distinct from the vowels *i* and *u*, and indeed, glides occur only in syllable onsets.¹³ We can thus assume that word-final glides are always incorporated into the word-final syllable rime as high vowels, leaving them beyond the scope of classification. The fact that *y* and *w* are not referred to by any Abu' class assignment rule is thus to be expected, given the phonotactic constraints on the distribution of these sounds.

The phoneme *k^w* is entirely absent from word-final position in Abu'. It occurs word-internally in the class XVII noun plural and plural numeral suffix (as in *dabak^{wih} bie-k^{wih}* 'fingers.XVII.PL two-XVII.PL'); otherwise, *k^w* only appears word-initially (*k^wapita* 'spoon', *k^wa'araka* 'frog'). Gaps in the distribution of *k^w* seem to reflect historical changes in the realization of this phoneme; labialized velars are cross-linguistically unstable and are often subject to simplification and loss over time, particularly in phonetically weak environments.¹⁴ In Abu' there is a set of nouns that require verbal agreement marked with *k^w-*; the nouns themselves, however, end in a glottal

¹²As it happens, no Arapesh dialect has the voiced coronal obstruent *d* word-finally. This distributional generalization holds even in dialects where other distinctively voiced obstruents are allowed in this position. I have no explanation for this pattern.

¹³One source of surface onset glides is a phonetic rule that freely applies within high-low vowel sequences to separate underlying adjacent vowels; e.g., *iah* [iyah] 'road'.

¹⁴Proto-Indo-European **k^w*, for example, simplified to plain *k* in Sanskrit (PIE **wlk^wos* 'wolf', **k^wri* 'buy' > Sanskrit *vrka*, *kri*); elsewhere **k^w* lost its primary place, palatalizing to *t* before front vowels in Greek (PIE **k^we* 'and' > Greek *te*; see Hock 1991:135).

stop, and the adjective agreement suffix corresponding to such nouns echoes the glottal stop, as in (8).

- (8) *numata'* *afu-'i* *k^w-ahe'*
 woman.II.SG good-II.SG II.SG-went
 'the good woman went'

Many glottal-stop-final nouns in Abu' have cognates in the other Arapesh dialects. As in Abu' they systematically belong to the feminine gender corresponding to Abu' class II.

- | | | | | |
|-----|-----------------------------|-------------------------|----------------|----------------|
| (9) | <u>Singular</u> | <u>Plural</u> | <u>Variety</u> | |
| | <i>aha'</i> | <i>ahawa</i> | Abu' | 'coconut palm' |
| | <i>afa'w</i> | <i>afawa</i> | Mufian | |
| | <i>ohok</i> | <i>ehemeb</i> | Bukiyip | |
| | <i>ahwok^u</i> | <i>ahwameb, ahwarib</i> | Cemaun | |
| | <i>numata'</i> | <i>numatawa</i> | Abu' | 'woman' |
| | <i>nemata'w</i> | <i>nematawa</i> | Mufian | |
| | <i>almatok</i> | <i>almagou</i> | Bukiyip | |
| | <i>aramatok^u</i> | <i>aramagou</i> | Cemaun | |

Given these correspondences, coupled with the fact that *k^w* does not occur word-finally in Abu', it is reasonable to assume that at least some, if not all, noun-final glottal stops in Abu' represent a debuccalization of *k^w* in final position.¹⁵ The impossibility of a class-assignment rule distinguishing noun-final (labialized) velars from glottal stop follows directly.

The absence of a gender rule referring to *r* in Abu' requires a similarly broad perspective to understand. For Papuan languages which have a distinct liquid, 'generally [r] and [l] are free or conditioned variants of... one phoneme' (Foley 1986:55; see also Wurm, Laycock and Voorhoeve 1975). In Arapesh, *r* and *l* phones both occur, though Fortune 1942:2 makes it clear that the difference between them is not contrastive. While Bukiyip does have both *l* and *r* phonemes, the phonological distribution of *r* is limited in that it does not occur word-finally (see Conrad and Wogiga 1991:2). In Mufian there is only one liquid (see Alungum, Conrad, and Lukas 1978:90), consistently recorded as *l*. In no Arapesh dialect are there distinct class-assignment rules referring to the two liquids.

¹⁵Simplification of final *k^w* is motivated by a general Arapesh rule that labialization is articulated as a labial offglide except when the preceding vowel is round; we thus get Cemaun [merik^w] *merik^w* 'rattan' but [oh^wok] *ohok^w* 'coconut palm'.

Nekitel states that in Abu' *r* is realized as a trill in word-final position, though he provides only one example of a final *r* on a native Abu' word, *bur* 'breakage' (the other *r*-final examples are borrowings: *pater* 'priest', *sister* 'sister'). The noun *bur* forms a minimal pair with *bul* 'pig', one of the most frequently occurring lexical items in the language; *bul* is also anomalous in that the noun-final liquid is preserved in the plural, though there it is realized as *r*: *burkuh* 'pigs'. Apart from this example, word-final liquids are consistently recorded as *l* (e.g., *alimil* 'bird', *wabul* 'village', and borrowed *raipel* TP 'rifle'). Native Abu' nouns systematically correspond to nouns in the *l*-final and *bVl*-final classes of Bukiyip and Mufian, and to nouns in the *r*-final and *bør*-final classes of Rohwim:

(11)			<u>Abu'</u>	<u>Rohwim</u>	<u>Bukiyip</u>	<u>Mufian</u>
	'bird'	SG	<i>alimil</i>	<i>aramir</i>	<i>almil</i>	<i>amil</i>
		PL	<i>alimikuh</i>	<i>aramigu^h</i>	<i>almiguh</i>	<i>aminguf</i>
	'village'	SG	<i>wabul</i>	<i>wabør</i>	<i>wabəl</i>	<i>wambel</i>
		PL	<i>walub</i>	<i>waryb</i>	<i>walúb</i>	<i>walemb</i>

Thus, while the *r* phone occurs with some frequency in Abu', its contrastive value in word-final position is limited, and class assignment makes no use of it.¹⁶

I have tried to show that the absence of Abu' class assignment rules that refer to the consonants *d*, *y*, *w*, *k^w*, and *r* is a fact about Abu' morphology that can be attributed to constraints independently provided by the language's phonology. There is just one segment yet to discuss, which is conspicuously absent from the Abu' class-assignment system despite being phonologically acceptable in word-final position, namely *s*. This phoneme cannot be ruled out of noun-final position except on morphological grounds: it appears in final position on the plurals of nouns in several Abu'

¹⁶There are some Tok Pisin borrowings that end in *r*, such as *pater* 'priest'. These new noun-final *r*'s do have the ability to trigger agreement in at least some syntactic contexts; thus *pater ara* 'priest this-*r*' now exists alongside *aul ala* 'eel this-*l*' (Nekitel 1986:193). As the word-final contrast between *l* and *r* becomes more firmly established phonologically, both sounds become morphosyntactically visible.

On the other hand, the morphological equation of *l* and *r* is thoroughgoing in Cemaun, a dialect in which *r* rather than *l* is the regularly occurring phone. Some borrowed *l*-final nouns are interpreted morphologically as if they ended with *r*. So the nouns *botol* 'bottle', *mabol* 'marble', and *bal* 'ball' are assigned both the canonical plural for native *r*-final nouns, *-guh* (*botoguh*, *mabologuh*, and *baloguh*, respectively), and agreement with the *r* paradigm (as in *botol cikin-i-r-um əbər* 'bottle full-POSS-*r*-OBL water', where the element *r* marks agreement with *botol*).

classes (e.g., *baraka* ~ *barakas* ‘head’ ~ ‘heads’, *aulaf* ~ *aulas* ‘house’ ~ ‘houses’), and it appears as the final element in one class of (singular) diminutives as well (e.g., *raitawas* ‘small portion of rice’, *lehitawas* ‘small portion of sago’). Thus, *s* is the lone counterexample to the generalization that the Abu’ classes are based on the dialect’s word-final consonantism. The special status of this segment is also noteworthy because in Mufian, Bukiyip, and Cemaun noun-final *s* has developed a new paradigm of its own exclusively for *s*-final borrowings which marks agreement with *s*. This development can be understood as filling a gap left open by the class-assignment morphology of these dialects. The innovation of an *s*-class is discussed in Chapter 5.

In sum, class assignment in Abu’ makes nearly exhaustive reference to the language’s word-final consonantal inventory. But as mentioned above, there are more classes in Abu’ than there are assignment rules referring to noun-final consonants. This can be accounted for by recognizing the coexistence of partially overlapping phonologically-based, morphologically-based, and semantically-based class assignment rules. An Abu’ noun ending in *n*, for example, belongs to masculine class 1 if the noun denotes a human male. If the final *n* is part of the diminutive suffix *-ikin* (e.g., *alemanikin* ‘small man’, cf. *aleman* ‘man’), the word belongs to the class specific to diminutive nouns. The remainder, *n*-final nouns that are neither diminutive nor masculine, are assigned to a uniquely *n*-final class. That the exclusively form-based *n*-final class is a productive category in its own right is supported by the assignment of *n*-final loanwords to this class. For example, the borrowing *baten* ‘button’ takes the same set of plural and agreement forms as the native Abu’ *n*-class noun *dubaun* ‘lobster, crayfish’. Hence, the claim that the Abu’ classes exhaust the language’s word-final consonantal phonology is made more precise by noting that class assignment does not stop at this point; rather, overlapping semantic cuts are made, as are finer formal cuts. For example, Abu’ has separate *bul*-final and *l*-final classes. The additional phonological information in *-bul* versus *-l* seems to constitute a case of proper inclusion, making assignment subject to the Elsewhere condition. One question we will be asking is whether the Elsewhere condition really does govern the application of class-assignment rules; if not, then some other means must be enlisted in order to assign a noun like *nubul* ‘belly’ to the *bul* class and prevent its assignment to the *l* class.

Having established the general properties of Arapesh classification, we turn to the details of assignment in the Rohwim dialect of Mountain Arapesh described by Fortune 1942. We will see that in Rohwim, plural form provides an additional constraint on class assignment, creating an extensive set of morphological assignment rules superimposed upon phonological criteria of the sort discussed above.

Noun Classification in Rohwim

Arapesh

Inflectional classes have a life of their own, existing as independent parts of the grammatical engine.

—Mark Aronoff, *Morphology By Itself*

But if I am for myself alone, what am I?

—Rabbi Hillel, *Pirkei Avot* 1:14

2.0 Introduction

This chapter is concerned with noun classification in Rohwim, the variety of Mountain Arapesh described by Fortune 1942 and spoken in the present-day village of Woginara. This system has been analyzed by Aronoff 1994 as a case study in ‘morphology by itself’, Aronoff’s program for highlighting those areas of morphological patterning which are not reducible to either phonology or syntax. Aronoff’s discussion will serve as our starting point here both for introducing the facts of Arapesh noun classification in more detail and for considering the general question of how qualitatively different types of morphological information should be integrated in a model. I will show that construing morphology as a series of continuous levels mapping between morphosyntactic and phonological representations as Aronoff does is problematic, because from such a perspective noun classification rules rooted directly in phonological form are unexpected, whereas that seems to

be precisely the correct generalization for Arapesh.¹ Instead, I show that morphosyntactic class and inflectional realization are better recognized as categories of information associated with the nominal lexeme, ‘the fundamental element in the lexicon of a language’ (Matthews 1991:26). These categories of information may be related systematically to one another and to other aspects of a lexeme’s representation, including its phonological form. Chapter 3 then explores the different ways these categories of information are structured. Descriptions will be couched in a broadly realizational approach to morphological structure (see Stump 2001). Reference to ‘morphemes’ throughout should be understood as shorthand for schematic rules relating two or more forms, like $X \rightarrow Xy$.

2.1 The *Morphology By Itself* Model

The model of morphology developed in Mark Aronoff’s *Morphology By Itself* (1994) emphasizes those core facts about word structure that cannot be derived by morphophonological rules or predicted from syntactic principles. Though it often has partial extra-morphological associations, such arbitrary morphology can be assigned no extra-morphological function. Its role is simply to organize the formation of words in a language, and supply some order to the association between abstract lexemes and their concrete phonological realizations. Aronoff calls such pure morphology ‘morphomic’ and treats it as belonging to its own grammatical level. The morphomic level is defined not by the distinctive nature of its input, nor by the specificity of its representational vocabulary, nor by the types of operations that apply within it. Rather, it is motivated by the existence of non-isomorphic mappings between morphosyntactic categories on the one hand, and their phonological realizations on the other. Aronoff illustrates this idea with the merging of two morphosyntactic categories, the passive and the past perfect, in a single formal realization in English, the perfect participle. The connection between the passive and the past perfect is by no means necessary in syntactic terms, and in many languages the two constructions are formally distinct. Yet their realization in English is always identical, no matter how they are formed (1).

¹Rules rooting noun class directly in phonological form are not expected in Aronoff’s view, though they are not strictly impossible. Aronoff posits one such rule but registers his discomfort with it by devoting to it a subsection entitled ‘A problematic case’ (1994:109-111).

- | | | | |
|-----|---------------|----------------|--------------|
| (1) | <i>bake</i> | <i>baked</i> | -D (default) |
| | <i>create</i> | <i>created</i> | |
| | <i>sing</i> | <i>sung</i> | ablaut |
| | <i>hang</i> | <i>hung</i> | |
| | <i>eat</i> | <i>eaten</i> | -N |
| | <i>be</i> | <i>been</i> | |
| | <i>feel</i> | <i>felt</i> | multiple |
| | <i>get</i> | <i>gotten</i> | |

The coherence of the perfect participle is thus attributable neither to its syntactic properties nor to its precise means of formation, since there are several of each. Instead, this significant though abstract aspect of English grammatical organization receives a unified representation at the morphomic level.

Among types of morphomic generalization is the inflectional class, which I will call here ‘realization class’: ‘a set of lexemes whose members each select the same set of inflectional realizations’ (Aronoff 1994:64).² Realization classes, the verbal conjugations and nominal declensions of classical grammar, are abstract from both phonological form and syntax in two important ways. First, realization classes involve *sets* of distinct inflections. In Latin, for example, each of five nominal declension classes provides inflectional endings for nouns in six cases in both the singular and the plural. The traditional analysis of the Latin first declension is given in (2). Since this entire series of phonological forms and morphosyntactic values is decided simultaneously, the unity of the paradigm must lie at some more abstract level of representation.

(2)

Case	Singular	Plural
NOM	<i>-a</i>	<i>-ae</i>
VOC		
ACC	<i>-am</i>	<i>-a:s</i>
GEN	<i>-ae</i>	<i>-a:rum</i>
DAT		<i>-i:s</i>
ABL		

²I prefer the term ‘realization class’ to ‘inflectional class’, as it avoids any implication of syntactic relevance along the lines of Anderson 1982, 1988, 1992.

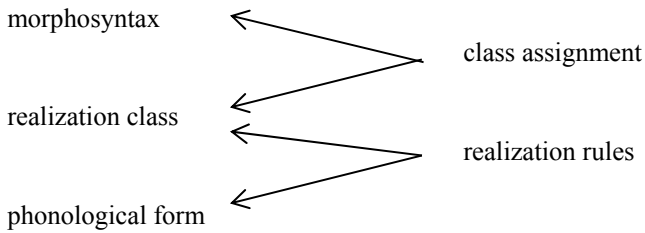
Realization classes are also abstract in that they are normally only imperfectly predictable from a lexeme's other properties, be they lexical properties such as semantics, stem form, theme vowel, or gender, or syntactically derived properties such as case. In other words, a lexeme's assignment to a realization class is not always derivable by rule, or even by a natural set of rules. Looking again to Latin for illustration, the fifth declension class is limited exclusively to nouns of masculine gender, and in the verbal morphology the first conjugation is regularly associated with the theme vowel *a*: (at least for imperfective stems). But neither the association between class and gender, nor that between class and theme vowel, holds throughout the language. Realization class is thus uniquely associated neither with morphosyntactic properties, nor with phonological form.

In nominal morphology, realization class must be distinguished from morphosyntactic noun class or gender. Like realization classes, noun classes divide the lexemes of a language into groups for purposes of determining inflection. But unlike realization classes, which determine the inflectional properties of noun themselves, noun classes determine the inflectional properties of associated words. Noun class or gender systems always show some degree of correlation with semantic features such as sex, animacy, humanness, or shape, and may also be associated with the formal properties of nouns. Nevertheless, such associations are rarely thoroughgoing within a system, and an individual item's class can only be definitively established by considering the inflectional properties of words to which the noun stands in a syntactic or anaphoric agreement relation.³

In Aronoff's model these two related but distinct types of morphological phenomena are conceptualized as distinct levels of nominal morphology, with their properties following from the different ways they interface with other levels of grammar. Noun class or gender often has semantic associations and is crucially syntactic in that it requires agreement to be marked on structurally associated words. Realization class is confined to the word itself, having consequences only for the noun's own phonological form. The levels model is schematized in (3) below, adapted from Aronoff 1994:71.

³Following Corbett 1991, 2006 I include the control of anaphoric pronouns by their antecedents within the scope of gender agreement. Pronominal and purely syntactic agreement targets like modifiers typically involve the same lexical categorizations in a language, and the two are linked in a single implicational 'agreement hierarchy' of considerable robustness. This assumption about pronoun agreement has consequences for the analysis of alliterative concord as discussed in Chapter 6.

(3)



Morphosyntactic agreement and phonological form are thus related indirectly through a distinct intermediate level that is identified with pure morphology. The model does not strictly exclude direct mappings between discontinuous levels; the intervening level could be transparent and so isomorphic with an adjacent level. Nevertheless, there is no reason to expect such discontinuous mappings to be regularly exploited by a grammar. Should we find that they are, the intermediate level's crucial role would be called into question, and with it the notion of hierarchical levels as the appropriate way to model the pattern.

Let us turn now to consider the respective roles of phonological form, purely morphological realization class, and morphosyntactic agreement in the Rohwim dialect of Mountain Arapesh. While these three properties are demonstrably distinct, they are nevertheless associated with one another in ways that are unexpected on a model in which morphology is construed as a level mediating the mapping between syntax and phonological form.

2.2 Arapesh Nouns

There is extremely limited morphological case marking in Rohwim, most participant roles being expressed by the order of NPs and class-marking affixes with respect to the verb. The unmarked word order in Arapesh clauses is SVO when the subject and object arguments are full NPs, although the fronting of non-subject arguments is often used as a strategy for topicalization. On verbs, the placement of class markers is a lexical property of the verb: some transitive verbs mark objects by way of prefixes to the root, others by suffixes. Both patterns are illustrated in the serial verb construction presented in (4). Intransitives are similarly split according to whether their one participant is marked with a prefix or suffix.

- (4) *fa-gw-a'* *fa-it-og^u*
 they-III.SG-eat they-finish-III.SG
 'they ate it all up' (referencing the class III noun *yabig^u* 'soup')

There is some formal differentiation among the pronominal affixes, particularly in the plural. Whereas in prefix position they tend to echo only the least sonorous segment of the noun terminal, in suffix position they may be expressed by a full CVC syllable. The class II noun *abør* 'water', for example, is referred to with the verb prefix *b-* in (5a), but with the suffix *-bør*—copying the entire final syllable of the noun—in (5b).

- (5) a. *abør* *ba-tagur* *ferororororo*
 water.II.SG II.SG-emerge 'gurgle-gurgle'
 'the water came gurgling out'
- b. *abør* *fa-bør-ah*
 water.II.SG they-II.SG-drink
 'they drank the water'

Nearly all nouns have distinct singular and plural forms. Which plural a given noun selects depends on both the singular form of the noun and the noun's realization class. The frame sentences in (6) and (7) illustrate two noun classes, I and IV, which correspond to three plural realization classes. In each variant of the sentence 'His white X fell down', the possessive pronoun 'his', the attributive adjective 'clean, white', and the verb 'fall down' all agree with the subject head noun in class and number.⁴

(6) Class I

<i>anan-i-b</i>	<i>bagab^v</i>	<i>bagara-bi</i>	<i>b-atu</i>
he-POSS-I.SG	floor.plank.I.SG	white-I.SG	I.SG-fell.down
<i>anan-i-bys</i>	<i>bagabys</i>	<i>bagara-bysi</i>	<i>s-atu</i>
he-POSS-I.PL	floor.plank.I.PL	white-I.PL	I.PL-fell.down

⁴A phonological characteristic of Arapesh that we first encounter in these examples is what Fortune calls the 'whispered terminals', colored releases on labial and velar stops in word-final position. Fortune does not usually transcribe the whispered terminals in his 'partly phonemic, partly phonetic' orthography, but I restore them here whenever they are predictable from his description. Fortune often transcribes final *h* and *hw* as ^h and ^{hw}.

(7) Class IV

<i>anan-i-k^u</i>	<i>arapefik^u</i>	<i>bagaro-kwi</i>	<i>kw-atu</i>
he-POSS-IV.SG	fem.friend.IV.SG	IV.SG-white	IV.SG-fell.down
<i>anan-i-u</i>	<i>arapefijer</i>	<i>bagara-ui</i>	<i>w-atu</i>
he-POSS-IV.PL	fem.friends.IV.PL _a	IV.PL-white	IV.PL-fell.down
<i>anan-i-k^u</i>	<i>yahak^u</i>	<i>bagaro-kwi</i>	<i>kw-atu</i>
he-POSS-IV.SG	fruit.tree.IV.SG	IV.SG-white	IV.SG-fell.down
<i>anan-i-u</i>	<i>yaharib</i>	<i>bagara-ui</i>	<i>w-atu</i>
he-POSS-IV.PL	fruit.tree.IV.PL _b	IV.PL-white	IV.PL-fell.down

Class I nouns do not have alternative plurals; they all take the same suffix *-bys*. Moreover, nouns of no other class regularly select this plural form. This means that for class I nouns, class and plural realization are isomorphic. For class IV nouns, by contrast, there is a set of phonologically irreducible allomorphs that may be used to mark the plural. The noun *arapefik^u* selects the suffix *-ijer* in the plural, while the noun *yahak^u* selects the suffix *-rib*. Which plural allomorph a given class IV noun selects is an arbitrary feature of that noun, though the arbitrariness is constrained, since nearly all nouns in the class mark plural in one of these ways, and nouns in no other classes do. It is this partially arbitrary selection for the form of the plural marker that constitutes an Arapesh noun's realization class. The distinction between morphosyntactic noun class and plural realization class is diagnosed by agreement. In (7) we see only one pattern of plural agreement for class IV, regardless of how plurality is realized on the noun.

An impressive array of elements require agreement for class and number in Arapesh. Cumulative class/number markers are used for cross-reference on adjectives, verbs (for subject, object, and oblique arguments), numerals, possessors (which agree with the possessed NP),⁵ independent and interrogative pronouns, and three series of demonstratives ('this near me', 'that near you', 'that over there'—i.e., 'near him/her/it').⁶ Fortune lists thirteen sets of agreement forms, one singular/plural pair corresponding to each of the thirteen Rohwim classes. Thirteen-way agreement occurs in the third person only; first and second person pronouns do not register class. The adjective and independent pronoun agreement paradigms are illustrated in (8).

⁵There is no possessive agreement in Abu'.

⁶There are other minor agreement patterns as well, e.g., adverbs may be marked for class and number in constructions akin to prepositional phrases: *arakohon-i-h-um wano^{uh}* 'middle-POSS-XII.SG-OBL battle.XII.SG' ('the middle of the battle').

(8) <u>Class</u>	<u>Adjective ‘white’ SG ~ PL</u>	<u>Pronoun SG ~ PL</u>
I	<i>bagara-bi ~ bagara-bysi</i>	<i>abab ~ ababys</i>
II	<i>bagara-børi ~ bagara-røbi</i>	<i>ababør ~ abarøb</i>
III	<i>bagara-gi ~ bagara-gasi</i>	<i>agag ~ agagas</i>
IV	<i>bagaro-kwi ~ bagara-ui</i>	<i>akwok^u ~ awau</i>
V	<i>bagara-mi ~ bagare-ipi</i>	<i>amum ~ apeipiⁱ</i>
VI	<i>bagara-ni ~ bagara-bi</i>	<i>anan ~ abab</i>
VII	<i>bagara-ni ~ bagara-mi</i>	<i>anan ~ amum</i>
VIII	<i>bagare-ñi ~ bagare-fi</i>	<i>eñeñ ~ efef</i>
IX	<i>bagara-pi ~ bagara-si</i>	<i>apapu ~ asas</i>
X	<i>bagara-ri ~ bagara-guhi</i>	<i>arar ~ agwagu^h</i>
XI	<i>bagara-ti ~ bagara-gwi</i>	<i>atat ~ agwag^u</i>
XII	<i>bagaro-whi ~ bagara-ruhi</i>	<i>awhoh ~ aharuh</i>
XIII	<i>bagara-hi ~ bagara-hi</i>	<i>ahah ~ aheh</i>

It is the nouns themselves that I am concerned with in this chapter; the agreement patterns are of interest mainly insofar as they can be used to diagnose class. As the list of nouns in (9) shows, number marking depends in part on a noun’s morphosyntactic class and in part on its plural realization. The noun class categories are distinguished by the roman numerals on the left, and morphological realization classes are distinguished by the canonical pairings of noun-final segments next to them. It should be clear that the system of classification groups nouns on the basis of their final segments, and that there is virtually no overlap between classes in this regard. I am assuming for the moment that singular nouns are morphologically simple, i.e., that the noun-final segments that are predictive of gender do not constitute singular morphemes or inflectional ‘word-markers’ of some kind (cf., e.g., Harris’s 1991 treatment of Spanish nouns). Extensive justification for this assumption is presented in Chapter 3.

(9) Rohwim Arapesh Noun Classes

<u>Class (SG ~ PL)</u>	<u>Singular</u>	<u>Plural</u>		
I <i>b^y ~ bys</i>	<i>ahoryb^y</i>	<i>ahorybys</i>	‘knee’	
	<i>bøb^y</i>	<i>bøbys</i>	‘betel nut’	
	<i>wab^y</i>	<i>wabys</i>	‘night’	
	<i>maryb^y</i>	<i>marybys</i>	‘hornet’	
II <i>bør ~ ryb</i>	<i>kaligebør</i>	<i>kaligerøb</i>	‘spider web’	
	<i>wabør</i>	<i>waryb</i>	‘village’	
	<i>ñemabør</i>	<i>ñemarøb</i>	‘string’	
	<i>arukwibør</i>	<i>arukwiryb</i>	‘dry coconut leaf’, ‘torch’	
III <i>g ~ (ga)s</i>	<i>dzebiwag</i>	<i>dzebiwas</i>	‘space under house on stilts’	
	<i>aijag</i>	<i>aijas</i>	‘leg’	
	<i>arukweg</i>	<i>arukwegas</i>	‘lobster’	
	<i>nubarig</i>	<i>nubarigas</i>	‘garden’	
IV <i>k^u ~ meb</i>	<i>tagiruk^u</i>	<i>tagirumeb</i>	‘species of bird’	
	<i>pugak^u</i>	<i>pugameb</i>	‘bonneting thatch’	
	<i>k^u ~ u</i>	<i>malik^u</i>	<i>maliu</i>	‘rattan species’
		<i>amagok^u</i>	<i>amagou</i>	‘fly’
	<i>k^u ~ rib</i>	<i>yahak^u</i>	<i>yaharib</i>	‘fruit tree’
		<i>awhok^u</i>	<i>awharib</i>	‘coconut palm’
	<i>k^u ~ ib</i>	<i>seraifuk^u</i>	<i>seraifib</i>	‘turtle’
		<i>unuk^u</i>	<i>unib</i>	‘star’, ‘ogress’, ‘anus’
	<i>k^u ~ guhijer</i>	<i>mabitek^u</i>	<i>mabiteguhijer</i>	‘duck’
		<i>ramahek^u</i>	<i>ramaheguhijer</i>	‘large ruffed lizard’
	<i>k^u ~ ijer</i>	<i>arapeñik^u</i>	<i>arapestjer</i>	‘female friend’
		<i>uruwhik^u</i>	<i>uruwhijer</i>	‘side post of a house’
	<i>k^u ~ omi</i>	<i>babwek^u</i>	<i>babwekomi</i>	‘grandmother’
		<i>jamek^u</i>	<i>jamekomi</i>	‘mother’
<i>k^u ~ heu</i>	<i>nigauwik^u</i>	<i>nigaliheu</i>	‘daughter’, ‘dau.-in-law’	
	<i>meganukik^u</i>	<i>meganubiheu</i>	‘sister-in-law’	

(9) Rohwim Arapesh Noun Classes (cont.)

	<u>Class (SG ~ PL)</u>	<u>Singular</u>	<u>Plural</u>	
V	<i>Vm ~ (e)ipi</i>	<i>abum</i>	<i>abeipi</i>	‘inner core of breadfruit’
		<i>daudam</i>	<i>daudeipi</i>	‘spider’
		<i>irum</i>	<i>iripi</i>	‘breadfruit’
		<i>warum</i>	<i>waripi</i>	‘sprouting coconut’
VI	<i>n ~ b</i>	<i>serugwan</i>	<i>serugwab</i>	‘fossilized wood’
		<i>femaun</i>	<i>femaub</i>	‘dugong’
	<i>n ~ ab</i>	<i>uman</i>	<i>umanab</i>	‘type of fruit’
		<i>lawan</i>	<i>lawanab</i>	‘tree snake’
VII	<i>n ~ m</i>	<i>afuken</i>	<i>afukenim</i>	‘older brother to man’
		<i>megan</i>	<i>meganomwi</i>	‘brother-in-law to man’
	infixing:	<i>niganin</i>	<i>nigamin</i>	‘son’, ‘son-in-law’
		<i>arapeñin</i>	<i>arapeñin</i>	‘male friend’
	suppletive:	<i>raminen</i>	<i>raheim</i>	‘husband’
		<i>awanin</i>	<i>arahim</i>	‘younger brother to man’
	derivational:	<i>mufupimin</i>	<i>mufuhem</i>	‘man from Mushu Island’
		<i>walipepimin</i>	<i>walipem</i>	‘man from Wallis Island’
VIII	<i>ñ ~ f</i>	<i>abotiñ</i>	<i>abotif</i>	‘long yam’
		<i>kaiñ</i>	<i>kai f</i>	‘bow’
	<i>V ~ has</i>	<i>bagi</i>	<i>bagihas</i>	‘variety of taro’
		<i>sumo</i>	<i>sumohas</i>	‘flying fox’
		<i>rau</i>	<i>rauhas</i>	‘house posts’
IX	<i>p^u ~ gwis</i>	<i>ilup^u</i>	<i>ilugwis</i>	‘feast’
		<i>wabigep^u</i>	<i>wabigegwis</i>	‘evening’
	<i>p^u ~ s</i>	<i>anip^u</i>	<i>anis</i>	‘bamboo water container’
		<i>fup^u</i>	<i>fus</i>	‘leaf’

(9) Rohwim Arapesh Noun Classes (cont.)

Class (SG ~ PL)	Singular	Plural		
X <i>r ~ gu^h</i>	<i>awhijar</i>	<i>awhijagu^h</i>	‘bat’	
	<i>jur</i>	<i>jugu^h</i>	‘snake’	
	<i>tutukwir</i>	<i>tutuwkigu^h</i>	‘cannibalistic ogre’	
	<i>autar</i>	<i>autogu^h</i>	‘testicle’	
XI <i>t ~ g^u</i>	<i>nybat</i>	<i>nybag^u</i>	‘dog’	
	<i>unut</i>	<i>unug^u</i>	‘bundle’	
	<i>t ~ tog^u</i>	<i>karaut</i>	<i>karautog^u</i>	‘treetop’
	<i>nokwat</i>	<i>nokwatog^u</i>	‘mouth’	
XII <i>uh~uh ~ ruh</i>	<i>bawe^{uh}</i>	<i>baweruh</i>	‘mountain road’	
	<i>ñuh</i>	<i>ñuruh</i>	‘vulva’	
	<i>uh ~ gwiruh</i>	<i>natage^{uh}</i>	<i>natagegwiruh</i>	‘red eyed lizard’
		<i>wano^{uh}</i>	<i>wanogwiruh</i>	‘ambush’
XIII <i>a^h ~ e^h</i>	<i>bara^h</i>	<i>bare^h</i>	‘bag for carrying infant’	
	<i>ha^h</i>	<i>he^h</i>	‘finger’, ‘toe’	
	<i>uh ~ ih</i>	<i>wiru^h</i>	<i>wiri^h</i>	‘hand drum’
	<i>aru^h</i>	<i>ari^h</i>	‘species of reed’	

Some aspects of the above list call for further comment. Fortune (1942:15-17) suggests that there are two distinct plurals for class III nouns, one replacing noun-final *g* with *s* (e.g., *baijag ~ baijas* ‘sago bark’), another retaining the *g* and adding *-as* (*abeg ~ abegas* ‘betel nut quid’). Unlike the plurals in most other noun classes, however, these two allomorphs are predictable by rule, specifically by the morphophonological rule block in (10). Class III nouns that end in *jag* or *w(h)ag* systematically receive the replacive plural; otherwise, they receive the additive version.⁷

- (10) 1. $\frac{1}{X} \left\{ \begin{array}{l} j \\ w(h) \end{array} \right\} ag/ \rightarrow \frac{1}{X} \left\{ \begin{array}{l} j \\ w(h) \end{array} \right\} as/_{[+PL]}$
2. $\frac{1}{X} g/ \rightarrow \frac{1}{X} gas/_{[+PL]}$

⁷The class III plurals have a similar distribution in Cemaun, although there are a few exceptions and flexible nouns that can take either plural (e.g. SG *mahog* ~ PL *mahos/mahogas* ‘variety of cucumber’). The choice between additive and replacive plurals is predictable in the western Arapesh languages, where the original noncontrastive final vowels are retained.

Similarly predictable are the two forms of singular nouns in class XII, again because their distribution depends on the phonological environment. The singular noun terminal *uh* is described by Fortune as ‘fully sounded’ following a consonant. After a vowel, Fortune transcribes it as *uh*, implying a lack of voicing. (In Cemaun, all class XII singulars simply end in the labialized glottal fricative *hʷ*.) Whereas class XII singulars of both of these sorts take the short *ruh* plural, in Fortune’s data *gwiruh* is selected only by nouns with the voiceless offglide. Thus, distinct realization classes must only be posited for a phonologically-defined subset of class XII nouns, namely those that end in a voiceless offglide. Nevertheless, for these nouns (and they comprise the majority), a realization-class distinction does exist.

Aronoff 1994 suggests that the two realization classes of class VI, the first replacing final *n* with *b*, the second adding the suffix *-ab*, can be reduced to one on phonological grounds, since ‘all of the items in the first subclass have *u* before the final *n*, while none of those in the second subclass do’ (1994:104). But in fact there are nouns in the first realization class that have not *u* but *a* in the critical position (e.g., *serugwan* ~ *serugwab* ‘fossilized wood’), and there is moreover a phonologically similar pair, contrasting in realization class, that has *i* in this position: *mañin* ~ *mafəb* ‘pigeon’ (first subclass), vs. *jehin* ~ *jehinab* ‘tree kangaroo’ (second subclass).⁸ I will thus keep close to Fortune’s description and assume that the two class VI plurals are morphologically distinct.

Finally, the distribution of the two class XIII plurals can be stated in phonological terms: the plural form maintains the height of the noun’s last vowel, changing it to its front counterpart in the vowel system; thus *a* → *e*, *u* → *i*. This pattern of fronting irrespective of vowel height holds for cognate plurals in other Arapesh varieties, e.g. Cemaun *yah* ~ *yeh/yegwih* ‘road’, *wirih* ~ *wirih* ‘hand drum, *witəh* ~ *wigweh* ‘gourd lime container’. Thus, although Fortune describes class XIII as comprising distinct subclasses, they should actually be subsumed under a single rule: /XV_[-FR]h/ → /XV_[+FR]h/[+PL]

Fortune states firmly that the Arapesh ‘system of noun classes is not a division of meaningful objects on any recognized principle of meaning, or of the form of things meant’ (1942:11). However, there is some semantic clustering within several of the classes, for example in classes V and IX, which include among their otherwise semantically diverse membership a large number of botanical terms. Fortune (1942:43-45) notes that in the botanical vocabulary there is a strong tendency for nouns referring to particular species of fruits and leaves to be ‘pulled’ into a quasi-paradigmatic clus-

⁸The stem-internal alternation *ñ* ~ *f* in ‘pigeon’ is an independent process which will be discussed later on.

ter by the basic level nouns *sulukum* ~ *suluweip* ‘fruit’ and *fup* ~ *fus* ‘leaf’, which belong to class V and IX respectively (10).

(11)	<u>‘tree’ (SG ~ PL)</u>	<u>‘fruit’ (SG ~ PL)</u>	<u>‘leaf’ (SG ~ PL)</u>
	<i>agøp</i> ~ <i>agøs</i>	<i>agum</i> ~ <i>agwip</i>	<i>agøpip</i> ~ <i>agøsis</i>
	<i>iluh</i> ~ <i>ilih</i>	<i>ilum</i> ~ <i>ilip</i>	<i>iluhip</i> ~ <i>iluhis</i>

It is not clear whether this should be understood as a unified phenomenon insofar as the fruit terms seem to use the class marking segments with a derivational function, whereas the leaf terms apparently incorporate the tree nouns into a possessive construction like *agøp-i-p* agøp-POSS-IX.SG but ‘without the need for adding *fup* [class IX ‘leaf’] being felt’ (Fortune 1942:44).

Whatever the case, proposals have been put forward that the classes have significant semantic associations; for example, Mead 1947:183 suggests that ‘the [class I] termination *ibis*’ which indicates social groups hailing from a common place, is ‘the ending for the gender which contains clustering things like eggs and areca nuts.’ Conrad 1996 develops this idea further, identifying semantic correlates for nearly all of the Arapesh noun classes. A few of these are quite enticing. For example class II, which includes the names for ‘spider web’, ‘banana peel’, various vines, and ‘hair’, is proposed to be the class of things that cling strongly to something else. Other proposed class meanings are more tenuous, such as the grouping together of ‘head’, ‘space of ground under a house’, ‘lobster’, and one type of bamboo under a category for things that are used or observed every day.

Whatever descriptive utility these kinds of semantic factors might have, there is little doubt that they are ranked below form as the decisive criterion for classification, at least synchronically, given the way new vocabulary is incorporated into the system. Borrowed terms for inanimates are classified overwhelmingly according to their form. Thus in Cemaun, *cok* ‘chalk’ is treated as an ordinary *k^w*-final class IV noun, pluralized following the regular form-based system as *comeb* and assigned alliterative *k^w* ~ *w* agreement in consequence (the noun-final *k* is interpreted as labialized due to coarticulation with the preceding round vowel). Similarly, inanimate *n*-final *mo^{fin}* ‘machine’ is classified according to its form as class VI, and pluralized accordingly as *mo^{fin}əb*. At the same time, those borrowings that would be obvious candidates for treatment along the lines of the most compelling semantic classes do not behave as their semantics would predict. So, for example, TP *bol* ‘ball’ and *mabol* ‘marble’ receive regular class X agreement and plural morphology, rather than the markers and agreements that would affiliate them with the spherical class proposed by Conrad.

Semantics does seem to influence the classification of nonhuman borrowings in one situation: Tok Pisin borrowings sometimes have agreement properties that are transparently influenced by the corresponding native term. So, for example, in Cemaun *helikopta* ‘helicopter’ takes class IV agreement despite its vowel-final form due to its association with the native noun *karahok*^w ‘canoe, vehicle’; *plet* ‘plate, dish’ takes class XII agreement with *h*^w in light of its native counterpart *yaureh*^w (it can also take class XI agreement with *t* following its form).

Two of the Arapesh noun classes do have semantic correlates. Class IV includes all nouns for female persons and roles, though many non-female nouns fall into this category as well. Class VII is the one purely semantic class: it includes all and only nouns denoting male persons and roles. Singular nouns in class VII are formally indistinguishable from those in class VI, as are their agreement forms. But the two categories are distinct in their plural morphology as well as in their semantics, and I will treat them as separate for purposes of this analysis (the issue is discussed further below). Outside of these clearly circumscribed areas of the vocabulary, however, I will assume that the noun classes have no active synchronic basis in semantics or conceptual structure.

Class VIII serves as the system’s default class in at least two of the three capacities discussed by Zubin and Shimojo 1993 for classifiers: it is used to classify nouns with unspecified referents, and it supplies a remainder category for nouns that fall outside the domain of other classes in the inventory. (It also occasionally serves in Zubin and Shimojo’s ‘default’ function as a substitute for a more specifically appropriate class). When for some reason the class of an NP cannot be determined in a given syntactic context, class VIII agreement forms are called upon to fulfill the morphological requirement that class be marked on syntactically associated elements. Class VIII agreement appears in default where the class of a noun phrase is inconclusive or unavailable, such as when the noun phrase lacks a specified head, as in the questions in (12), or in pronominal reference ‘where the class of the noun specified is not specially singled out’ (Fortune 1942:49; see also Aronoff 1994:97-103).

- (12) a. *amwiña* *ñi-gure*
 who.VIII.SG VIII.SG-cry.out
 ‘Who cries out?’
- b. *ne-lik* *nabep* *neña-kali* *mane-f*
 VII.SG-ask nabep.bird 2.SG-like what-VIII.PL?
 ‘He asks the *nabep* bird, “what would you like?”’

In (13) we see an example of the typical Arapesh pattern of default agreement with conjoined noun phrases of conflicting classes. The conjuncts are the object noun ‘blackbird’ (which is elevated into the male human class VII due to its social role as a male in the legend from which the example is extracted) and blackbird’s wife, of class IV. Because the conjuncts differ in class, they make conflicting demands on the verb object marker. The conflict is resolved by suffixing the verb with a class VIII object marker in default.

- (13) *ni^h* *ha-gak-ef* *tabali*
 fire.XIII.PL XIII.PL-dies-VIII.PL blackbird
 anan *ni* *irohokwik^u*
 himself.VII.SG with wife.IV.SG

‘Fire kills blackbird and his wife.’

There are also some Arapesh nouns that are anomalous from the point of view of the predominantly form-based system, and these take class VIII agreement as well. There are at least four distinct types of anomalous nouns. The first includes formally regular but idiosyncratically ‘classless’ nouns (14); five clear examples are attested in Fortune’s texts. Each has the regular formal correlates—i.e., the realizational properties—of one of the thirteen systematic Arapesh classes, both in the singular and the plural. Yet nouns belonging to this group do not pattern in agreement along the lines expected on the basis of their form. Instead, they take class VIII agreement, apparently in default, as seen in (15).

- (14) *diliat ~ diliatog^u* ‘side post’
 (*class XI)
lawag ~ lawas ‘tree’
 (*class III)
wehehag ~ wehehas ‘sago spike’
 (*class III)

- (15) *wo-no^h* *wehehas* *ebit*
 IV.PL-thrust.in sago.spike.VIII.PL penis
 wo-ñ-no^h *wa-tugwar-iñ*
 IV.PL-VIII.SG-thrust.in IV.PL-snap.off-VIII.SG

‘They thrust sago spikes into his penis; they each thrust one in and snapped it off.’

The second type of exception includes examples from two formal subclasses. These two small groups of nouns consistently pair particular singular and plural forms, and so would seem to constitute minor realization classes in their own right. One group pairs the terminal sequence *gør* in the singular with *gu* in the plural, as shown in (16).⁹

- | | | |
|------|--------------------------------|--------------------|
| (16) | <i>natagør</i> ~ <i>natagu</i> | ‘cliff, precipice’ |
| | <i>ehigør</i> ~ <i>ehigu</i> | ‘sago palm trunk’ |

The other group, which has five members in Fortune’s text, pairs *n* with the *bys* plural of class I, as shown in (17).

- | | | |
|------|---------------------------------|----------------------------|
| (17) | <i>ahøn</i> ~ <i>ahøbys</i> | ‘grass rats’ |
| | <i>gusian</i> ~ <i>gusiabys</i> | ‘edible species of lizard’ |

Nouns in these subclasses do not receive agreement according to any of the regular Arapesh gender patterns; nor do they take agreement markers that match their phonological form. Instead, they take class VIII agreement, again, apparently in default.

The other two types of anomalous nouns are particularly interesting because they show how default class can be arrived at directly by way of the special plural ending *-ehas*. In Fortune’s materials I have found five native Arapesh nouns that take the *-ehas* plural because their final segments do not correspond to any specific plural, as in (18). Apparently, the atypical properties of these nouns prevent them from participating in regular form-based plural and class assignment.

- | | | | |
|------|---------------------------------------|----------------------|-----------------------------------|
| (18) | <i>bokok</i> ~ <i>bokokehas</i> | ‘cannibalistic ogre’ | (no <i>k</i> class) ¹⁰ |
| | <i>kwagesab</i> ~ <i>kwagesabehas</i> | ‘croton’ | (no <i>b</i> class) ¹¹ |
| | <i>pas</i> ~ <i>pasehas</i> | ‘taro pounder’ | (no <i>s</i> class) |

⁹The cognate class in Cemaun Arapesh, which also takes class VIII agreement, pairs singular nouns ending in *gir* with the class X plural *guh*.

¹⁰The noun *bokok* is a compound formed from two verb roots, *bo* ‘kill’ and *kok* ‘eat’. In Cemaun deverbal nouns (e.g. *dækem* ‘knowledge’, *hir* ‘lifting’) regularly take default agreement, irrespective of their phonological properties.

¹¹According to Fortune the final *b* in *kwagesab* and its sister exception *mib* ~ *mibehas* ‘thigh’ lack the rounded final release that occurs on class I nouns. I detect no such release on *b*-final nouns in any variety of Arapesh, and only possible remnants of a former labial release in agreement with some *p*-final singulars (e.g. *ñegwip at-up* ‘nipple.IX.SG one-IX.SG’ but *ñimiəməp at-ip* ‘species.of.ant.IX.SG one-IX.SG; cf. Fortune 1942:11-13). Class IX plural allomorphy is impervious to this distinction, with *-s* and *-gwis* occurring on nouns of both types.

Finally, Fortune's materials include eleven nouns that we would expect to belong to one of the systematic Arapesh classes given their phonological form in the singular, yet for some reason they do not. Instead they take *-ehas* in the plural and agreement with class VIII. Three examples are given in (19).

- (19) *sam* ~ *samehas* (**saip*¹) 'taro and coconut croquette'
iwan ~ *iwehas* (**iwab* or *iwanab*) 'arm ring'
tatuwar ~ *tatuwarehas* (**tatuwaguh*) 'arrow'

2.3 The Mapping Between Phonological Form, Morphological Realization, and Morphosyntactic Class

2.3.1 Aronoff's Analysis

The nouns taking the *-ehas* plural, particularly those of the type represented in (19) above, play an important role in Aronoff's analysis (1994:107):

[I]f form directly determined gender, independently of [plural realization] class, we would expect to have nouns whose gender corresponded to their form, even though they were exceptions to the independent [realization]-class-assignment rules. This is precisely the sort of nouns that we do not find.

In other words, a noun that takes the *-ehas* plural takes class VIII agreement, not only in the plural, but in the singular as well. Plural realization class thus plays a crucial mediating role in the mapping between phonological form and morphosyntactic agreement. Aronoff's analysis, which is designed to reflect this central role of realization class in the Arapesh system, follows a 'doubly inverted' prediction chain, whereby a noun's plural realization is a function of its phonological form in the singular, and a noun's morphosyntactic class depends in turn directly on its plural (Aronoff 1994:105). Aronoff calls this schema 'inverted' because both of its clauses are rooted in form.¹²

There are two main exceptions to the doubly inverted pattern. Class IV makes for something of a problem because it has at least eight plural realization patterns corresponding to only one singular form and noun class.¹³

¹²Aronoff 1994:74 considers this pattern of predictability 'peculiar' on the assumption that 'gender (being syntactic) is prior to [realization] class in the same way that syntax is prior to morphology'. But note that it is also exhibited systematically by the Russian declension system (Corbett 1991, Spencer 1994).

¹³The precise number of plural alternants is difficult to determine for classes IV and VII because they include so many kinship terms and other nouns referring to the social domain. This

Since k^u -final nouns map onto many plural realization classes but only one noun class (class IV), Aronoff posits ‘a direct mapping from form to gender for k^u nouns alone’ (1994:110), bypassing the level of realization class in this one instance. Class VII is a different story altogether in that this class of nouns has an exceptionless semantic basis, male humans, in addition to a formal one. ‘For this gender alone, it seems reasonable to assume that the conceptual system drives the morphology and the syntax quite directly’ (1994:112). So for class VII nouns we first have the meaning, then the form.

2.3.2 Subverting the ‘Double Inversion’

While Aronoff’s ‘doubly inverted’ prediction chain systematizes the Arapesh facts in a neat, intuitive way, it is unsatisfactory in two important respects. First, it fails to predict plural realization in a substantial number of cases. Second, it fails to express the fundamental generalization that so clearly drives the Arapesh system, namely that noun classes are rooted directly in singular-final phonological form. Addressing these problems requires a reevaluation of the default plural *-ehas*, the issue to which we now turn.

The reality of the Arapesh plural realization classes is that they are numerous and arbitrary. There are alternative plurals in nearly half the noun classes (classes IV, VI, VII, IX, XI and XII), rendering realization class unpredictable for a large proportion of Arapesh nouns. Consider, for example, the nouns *nybat* ‘dog’ and *alit* ‘shelf’ in (20). The choice of plural has no principled synchronic basis; which one a noun happens to take is simply a fact that must be memorized for each noun.

(20)	<u>Class X</u>	<u>Singular</u>	<u>Plural</u>	
	$t \sim g^u$	<i>nybat</i>	<i>nybag^u</i>	‘dog’
		<i>alit</i>	<i>alitog^u</i>	‘shelf’

At the same time, however, it is clear that the choices are limited, in the case of this class, to one of two alternatives. Such constraints on the form of plural marking hold for nearly all Arapesh nouns, because the realization classes do not normally overlap, applying instead only to nouns of a single class.¹⁴ So once a noun’s plural form is known its final segments in the singular can also be predicted. I am not suggesting that singular forms are lit-

area of the lexicon is often idiosyncratic, an interesting topic in its own right (see Dobrin 1997, Dahl and Koptjevskaja-Tamm 2001).

¹⁴The main exceptions to this statement are nouns that take the plural *-ehas*, to be discussed below, and the two groups of nouns shown in examples (12) and (13) above. These nouns require special marking under any analysis. See also note (22) below on the human plural *-omi*.

erally derived from their plural counterparts by way of a truncation process; the observation is rather that even when a plural has to be listed (because it is not predictable), its relation to a category of singular form is still at least partially regular. The predictive relation between realization class and phonological form is thus the reverse of what Aronoff's analysis suggests. The phonological form of a noun (i.e., the singular form) is predictable from its plural realization, rather than the other way around.

As for the morphosyntactic noun classes, a glance again at the data in (9) makes it clear that the classes overwhelmingly divide along lines distinguishing not the plurals, but rather the singular forms of nouns. This generalization is most obvious for class IV because it has an unusually large number of formally disparate plurals corresponding to a single form in the singular. But it is true to a lesser extent for nearly all the other classes as well. The classes that require some comment in light of this statement are (in order of increasing complexity) XIII, VIII and VI/VII. I will address them each in turn.

The problem presented by class XIII was already discussed above: the two singular-final forms Fortune identifies with this class, *a^h* and *u^h*, are reducible to a unitary phonological description, *V^h*.¹⁵ This fits in perfectly with the generalization that the classes are isomorphic with phonologically specifiable categories of singular-final form.

The two singular patterns attributed so far to class VIII, *ñ* and *V*, raise a question of a different sort: Are nouns ending in a voiced vowel assigned to class VIII by rule, or do they instead receive class VIII agreement by default? Vowel-final nouns are by no means uncommon: an exhaustive search of Fortune's texts uncovers thirty-two unambiguous examples, which is more than the total number of nouns listed for some classes (such as class II, which is small and evidently closed). There are grounds for classifying vowel-final nouns by default rather than by a rule referring specifically to their phonological form. For one thing, the terminal element *V* is at an unusually high level of phonological description; unlike the other phonological predictors it is at the broad level of major class, indifferent to the variations in vowel quality it subsumes. Assuming that noun-final vowels are outside the scope of the regular class-assignment rules also makes sense from a family-level perspective, since final vowels are systematically ignored by the class-assignment rules in other languages of the family such as Abu' as discussed in Chapter 1. In some cases it is possible to identify vowel-final class VIII nouns as borrowings from a neighboring language (e.g. *yawi* 'garden', *saki* 'myth, moral speech', or *mahi* 'species of frog' are from Boikin).

¹⁵In Cemaun the singular-final vowel is specifiable as -FRONT. This also appears to hold in Rohwim.

Finally, unlike the regular members of class VIII, which end in \tilde{n} in the singular and take agreement forms with \tilde{n} , elements marking agreement with vowel-final nouns do not echo the nouns' phonological form. Instead, they take agreement forms with \tilde{n} .¹⁶ So although they are numerous, nouns ending in a vowel are odd in enough ways to justify treating them as peripheral to the regular form-based class-assignment system, which means that they would receive class VIII agreement by default rather than by a rule referring specifically to their phonological form. (Possible reasons to revise this decision will be explored in Chapter 5.)

Classes VI and VII raise a different issue still, and probably the most troubling one for the phonological prediction scheme. Nouns in these two classes share a single phonological description in the singular; they both end in n , and they have indistinguishable agreement markers: all n -final nouns take agreement forms with n . It is thus not possible to predict gender on the basis of phonological form in the case of classes VI and VII. There are two different ways to try to resolve this problem. One would be to say that classes VI and VII are really not morphologically distinct in the singular. This solution would dissociate class from lexemes as such, so that 'class VI' and 'class VII' would only be meaningful for the plural forms of n -final nouns. The other approach, which is the one I will adopt here, relies on the fact that that n -final nouns belonging to class VII refer always to members of the category of human males, whereas n -final nouns with all other semantic associations fall into class VI. If semantically based assignment is assumed to take priority, as I will show in Chapter 4, the problem effectively disappears, since with the male human nouns diverted into class VII the remaining n -final nouns all shuffle neatly into class VI. This has the added advantage of keeping their plurals separate as well. So once provision is made for class VII, the generalization is robust: morphosyntactic noun class is regularly predictable directly from the final phonological segment of the singular form of nouns.

¹⁶There are two other situations in which agreement marking fails to formally echo the noun-final phonological element of a third person form: (1) in the plural with many of the class IV nouns, and (2) in the isolated case of the class VII plural subject prefix, which is $h-$, whereas the various class VII plurals (and all the other members of the class VII plural agreement paradigm) involve the segment m . But in these two cases the lack of alliteration occurs in the plural rather than the singular, whereas what I am trying to establish is the correspondence between gender and the phonological form of singular nouns. As regards class IV plurals, I argue in Chapter 4 that there is a phonological 'common denominator', a feature that defines the central tendency for plurals of this class, even though what they share cannot be reduced to it. This is the phonological feature +LABIAL, the feature that is distinctively present in the class IV agreement forms (w or u , depending on syllabic position).

The one situation in which class correlates specifically with something *other* than singular form is with nouns that take the plural *-ehas* (Cemaun *-əhəs*), ‘the diagnostic for membership in the default [realization] class’ (Aronoff 1994:105). In Fortune’s data, the *-ehas* plural unfailingly corresponds to class VIII, and this is clearly the crucial observation that prompts Aronoff to see the Arapesh class-assignment rules as ‘inverted’, depending directly on realization class rather than on phonological form. In what follows, however, I will argue that the correlation between *-ehas* and class VIII represents a significant departure from regular class- and plural-assignment patterns, warranting a separate account.

One approach to understanding the special behavior of *-ehas* is to explore its markedness or degree of irregularity relative to the other Arapesh plurals. My reasoning here is modeled after Janda 1990, which deals with a very similar problem in German plural formation. German also has a diverse set of plural markers, *-(e)n*, *-e*, *¨-e*, *-er*, *¨-er*, *-Ø*, *¨-Ø* and *-s*, one of which might be predicted to expand its domain of application at the expense of others over time (Janda predicts that German plural *-s* will spread). Appealing to markedness must be done with caution, however, since wildly different criteria that have been used to gauge the concept and not all are equally helpful (see Battistella 1990, Haspelmath 2006). For example, we know that even formally marginal patterns are resilient to change when they occur with high frequency, so that text frequency alone would be a deceptive indicator of a morphological element’s susceptibility to generalization (see Bauer 2001; cf. Greenberg 1966). But a conceptually distinct set of non-frequential criteria can highlight distributional asymmetries in a pattern, so following Janda 1990, I consider the Arapesh *-ehas* plural with respect to three conceptually distinct, non-frequential criteria for gauging morphological markedness: distributional breadth, redundant marking, and productivity.

Distributional breadth refers to the diversity of environments in which an element occurs. When there are ‘several competing markers which indicate the same morphological category’, all of which

retain some degree of productivity... subject to differing conditions on their contexts of occurrence..., it often turns out that one form occurs in a much more disparate set of contexts than the others and so can be considered as the ‘elsewhere’-case—i.e., as the unmarked default, given by a general rule. [Janda 1990:139-140]

As we have already seen, this is a way in which *-ehas* clearly stands apart as unmarked with respect to the other Arapesh plurals: *-ehas* can attach to nouns of any phonological shape, whereas virtually every other plural ending in the language is severely restricted in its formal domain of attachment.

But *-ehas* stands apart from the other plurals in apparently marked ways as well. First, *-ehas* is limited in its ability to serve as a redundant, double, or ‘pleonastic’ marker of plurality. The ability to serve as a secondary marker is useful indicator of markedness, because regular morphological elements are often extended to irregular forms (see Thomason 1988), while the reverse is less often found. An example is nonstandard English forms such as *stoled* and *broked*, where the regular past tense suffix *-D* is added to stems already irregularly marked for the morphological feature +PAST. Now in Arapesh there are a large number of nouns that are doubly marked in the plural, constituting a special group of what Fortune calls ‘intervocalic change’ (abbreviated here ‘IVΔ’) nouns. The class IV noun *mapok^u* ‘brown toad’, for example, takes the *-u* plural, which is formally appropriate for a noun ending in the class IV terminal *k^u*. Simultaneously, however, the word exhibits an internal alternation following another canonical pattern, namely the *p ~ s* pattern characteristic of class IX, so that the plural form is actually doubly marked *masou* rather than **mapou*. Further examples of intervocalic change are given in (21), where the parenthesized roman numerals represent the class typically associated with the intervocalic alternation when it occurs noun-finally. It is always the outer marker that determines agreement.

(21)	<u>Class (IVΔ)</u>	<u>Singular</u>	<u>Plural</u>	
	IV (IX)	<i>mapok^u</i>	<i>masou</i>	‘brown toad’
	VI (VIII)	<i>wiñun</i>	<i>wiføb</i>	‘patterned nut bag’
	VIII (VII)	<i>aramaniñ</i>	<i>aramumwif</i>	‘male infant’
	IX (XI)	<i>karuatepⁱ</i>	<i>karuagwes</i>	‘pigsticking spear’

We will have much more to say about these kinds of forms later, but for the moment the question is whether *-ehas* participates in intervocalic change as do the other plural formatives. The answer is essentially no. Unlike most plurals, which can occur outside of (i.e., in addition to) the plurals of other classes, *-ehas* appears as an unambiguous secondary plural marker on two Rohwim nouns, which are shown below in (22).

(22)	VIII (VIII)	<i>kaiñ</i>	<i>kaiſ(ehas)</i>	‘bow’
		<i>kobiñ</i>	<i>kobiſ(ehas)</i>	‘ditch’

These examples are unusual in that the secondary marker is not only optional but moreover added only to nouns already marked for the same class,

namely SG \tilde{n} ~ PL f .¹⁷ There is one example in Fortune's grammar which has *-ehas* appearing as the internal plural in an intervocalic change noun: *subariap^u* ~ *subarihasis* 'leaf of the *subari* tree', but as mentioned earlier, it is precisely these kinds of leaf and fruit terms that are also open to analysis as productive possessive formation rather than redundant plural inflection. So *-ehas* does not appear to be particularly unmarked according to the criterion of redundant marking (and this is true whether it is the external or the internal alternation that is held to be redundant).

But more compelling evidence for the non-typical status of *-ehas* comes from the treatment of Tok Pisin loan words in Arapesh.¹⁸ Borrowed nouns receive both their plurals and their patterns of agreement marking not by default, but instead according to the regular form-based patterns. The noun *blanket*, for example, is pluralized according to the canonical pattern $t \sim g^u$, and so its plural is *blanketog^u*, not **blanketehas*. The noun *blanket* moreover receives class XI agreement just like an ordinary Arapesh t -final noun, and does not take default agreement with class VIII. Some of the borrowings Fortune observed are listed in (23); an example from his texts showing regular agreement with a borrowed noun is given in (24).

(23)	<u>Class</u>	<u>Singular</u>	<u>Plural</u>	
	IV	<i>tommihawk^u</i>	<i>tommihaguhijer</i>	'tomahawk'
	VI	<i>sipun</i>	<i>sipunab</i>	'spoon'
	X	<i>neil</i>	<i>neiguh</i>	'nail'
	XI	<i>hat</i>	<i>hatog^u</i>	'hat'
(24)	<i>f-ek</i>	<i>anun (*eñeñ)</i>	<i>rain</i>	<i>na-pwe (*ña-pwe)</i>
	they-form	one.VI.SG	line.VI.SG	VI.SG-stay
	'They make one line'			

Only when a borrowed noun cannot be assigned a plural on the basis of its form does it take *-ehas* and class VIII agreement in default; examples are given in (25a). The sole exception cited by Fortune is the noun 'policeman', which is formed with the native Arapesh demonym *-pimin* (25b). This noun

¹⁷Fortune 1942:41 otherwise notes only 'a solitary example of the intervocalic change agreeing with the terminal', namely *jakerok^u* ~ *jaberemeb* 'tail of tree kangaroo'. Another example, *niganin* ~ *nigamim* 'son, son-in-law', has the alternative plural *nigamin*, with simple infixing.

¹⁸Fortune refers to these (and renders them orthographically) as English, but their immediate source is the English-based creole Tok Pisin. Tok Pisin was already being learned as a colonial language by the time of Fortune's fieldwork in 1931-32. Margaret Mead, with whom Fortune did much of his Arapesh fieldwork, in fact conducted the better part of her research on Arapesh culture in Tok Pisin (Mead 1940:337).

receives an *m* plural and the masculine class VII agreement expected in light of its male person semantics, showing that it is actually regular in the context of the overall system.

(25)	<u>Class</u>	<u>Singular</u>	<u>Plural</u>	
a.	VIII	<i>pusi</i>	<i>pusihas</i>	‘cat’
	VIII	<i>bullamakou</i>	<i>bullamakehas</i>	‘cow’
	VIII	<i>hors</i>	<i>horsihis</i>	‘horse’
b.	VII	<i>polisipepimin</i>	<i>polisihem</i>	‘policeman’

It is thus not the default plural *-ehas* but instead the entire system of form-dependent plural marking that extends productively to new nouns.

In sum, the three non-frequential markedness criteria we have considered give conflicting results. On the one hand, the *-ehas* plural is freer than other plurals in its domain of attachment. On the other hand, we might then expect it to more readily apply as a secondary marker of plurality on already plural nouns, and we might expect it to be the allomorph that is most regularly extended to new nouns, along the lines of unmarked English plural *-S*.

Before pausing to evaluate the distributional facts just described there is one more set of observations that can be brought to bear in determining the relationship of *-ehas* to other Arapesh plurals. Within each class, the alternate plurals do not simply pattern as random lists. Instead, the alternate endings tend to share formal features and so really constitute plural schemas structured in terms of family resemblances, or morphological constellations (Bybee and Moder 1983, Bybee 1988, Köpcke 1988, Janda and Joseph 1999). This is true even for the semantically motivated classes IV and VII, which have many formally diverse plurals: nearly every class IV plural includes the feature +LABIAL, and nearly every class VII plural includes the segment *m*. Furthermore, the plural alternants within a class do not share their domains equally. The frequencies in (26) show that for each class there is typically a main plural, one that appears on the majority of nouns with a given singular form, and one or more less common plural endings.¹⁹ As a

¹⁹These frequency counts are based on a survey of all nouns in Fortune 1942, both those listed in his grammatical description and those found in the 60 single-spaced pages of associated texts, which are assumed to be representative for the sake of this analysis. In most cases he states quite clearly that the distribution of the various plural realizations within each class is skewed. So for example Fortune says of class IX that ‘[t]he regular form is the change of the singular termination *p* to the plural termination *s*’, but only ‘[a] small subset drop the *p* of the singular termination and substitute *gwis* to form the plural’ (31); of class XI he says ‘[i]n the great majority of members of this class the plural is formed by the addition of *og^u* to the singu-

result, if an analysis is to treat any plural as unmarked, it should be the main form-based plural rather than the default plural *-ehas*.

(26) <u>Class</u>	<u>Plural Frequency</u>
III	<i>gas</i> (37), <i>s</i> (9)
IV	<i>meb</i> (14), <i>u</i> (11), <i>ijer</i> (7), <i>guhijer</i> (6), <i>rib</i> (4), <i>heu</i> (4), <i>omi</i> (3), <i>ib</i> (2)
VI	<i>b</i> (20), <i>ab</i> (4)
IX	<i>s</i> (45), <i>gwis</i> (6)
XI	<i>tog^u</i> (25), <i>g^u</i> (3)
XII	<i>ruh</i> (25), <i>gwiruh</i> (10) ²⁰

This whole array of facts makes sense if pluralization with *-ehas* is taken to represent a default entirely from the form-based system rather than the unmarked plural within that system. The limited use of *-ehas* in redundant marking and borrowings is evidence for the priority of the whole form-based assignment system, and its greater environmental diversity follows simply from its application to the residue, whatever nouns are left over once all the form-based assignments have been made.

According to this analysis (which will be useful as a starting point though we will consider other possibilities later on), a noun can default from the form-based system at either link in the [plural realization → singular form → noun class] prediction chain. A noun belonging to no canonical plural class will default at the first link; an example is the noun *lim* ~ *limehas* ‘roller for launching a canoe’. This noun violates the form-based schema by virtue of what I assume to be a specific marking for the *-ehas* plural, which does not have a predictable relationship to any canonical category of singular form. The noun’s class is therefore likewise arrived at by default (i.e., by following its plural).

The other possibility is for a noun to exhibit the expected relation between singular and plural form, and default in its link to noun class alone. Such a noun will have to be shunted out of the form-based system by an arbitrary marker in its lexical entry. Thus, for example, the exceptional noun *wehehag* ~ *wehehagas* ‘sago palm thorn, spike’ receives default class agreement despite the regular relation between its plural realization and singular form.

lar form’ (35), while there are only a ‘few members... [that] do not add *og^u*..., but drop the singular terminal *t* and substitute *g^u*’ (36).

²⁰The frequencies for class XII only take into account those nouns that Fortune transcribes as ending in *V^{uh}*, since it is only in this subset of class XII that plural realization classes are distinguished.

(29) Exceptional specification with *-ehas* (SG *lim* ~ PL *limehas*)

- a. Lexical representation: LIM ‘canoe roller’
*limehas*_[+PL]
- b. General form-based rule: $/Xm/ \rightarrow /Xipi/_{[+PL]}$
— blocked by the specific plural —
- c. Most general redundancy rule: $/X/ \leftrightarrow /Xehas/_{[+PL]}$

What is special about *-ehas* is that it is the only Arapesh plural not subject to phonological restrictions on its application. This property is expressed by the fact that *-ehas* is related to only a variable in (29c): the rule identifies *-ehas* as a plural marker, but is silent about the form to which it attaches. Given the *-ehas* plural’s ostensibly ‘Elsewhere’ distribution, it may perhaps seem counterintuitive that it is assumed here to be underlying in the lexical entries of exceptions like *lim*, and that *-ehas* is construed as a redundancy rule. It should be kept in mind, however, that something special must be said about these nouns in any case in order to block the more specific (but regular) rule (29b) from applying, and while it would be possible to posit an abstract diacritic on the noun to perform this blocking function, the main reason I can see for doing so would be to ensure that all applications of the *-ehas* rule are feature filling or ‘incremental’ in the terminology of Stump (2001). But it is not obvious that unifying the rule’s mode of application is necessarily desirable. After all, the dual nature of morphological rules—their ability to apply both as redundancy rules parsing existing forms and as productive rules creating new forms—has long been recognized within theories of word formation (see Jackendoff 1975, Aronoff 1976), and to the extent that rules of inflection exhibit varying degrees of productivity, such dual functioning seems to characterize inflectional morphology as well.

This is essentially the conclusion that was reached by Fraser and Corbett 1997 based on their work using the formal language DATR. To implement the Arapesh facts computationally, Fraser and Corbett find it necessary to recognize two distinct types of defaults, which they call ‘normal case’ defaults and ‘exceptional case’ defaults: A ‘normal case default is retrieved after failing to find any more specific value; an exceptional case default is retrieved after finding too much information—information which blocks normal retrieval’ (1997:44). While Fraser and Corbett’s lexical entry for the noun *lim* uses an exceptionality marker rather than the concrete plural form *-ehas* to do the blocking, they specifically note that ‘the account would still go through successfully if [the exceptionality marker] were re-

placed with an arbitrary symbol such as SOMETHING_FUNNY_HERE. In other words, the account is capturing the fact that anything which fails to fit the normal pattern ends up being assigned to the same class' (1997:40). In the analysis here, the 'something funny' is simply the *-ehas* plural itself.

Perhaps the most serious problem for our analysis is the fact that the majority plural is not always the productive plural, if productivity is gauged by application to borrowed words. The analysis here predicts Tok Pisin loans not only to receive the form-based plurals, but to be specifically subject to the majority plural rule, since that is the most general 'local default', the rule that applies when the appropriate contextual features are present and there is no more specific information in the lexical entry to block it. As discussed above, most loan words are adapted to Arapesh morphologically, receiving plurals that are appropriate given their phonological form. Some borrowed nouns do receive the local default. Borrowings ending in *t*, for example, systematically receive the additive plural ending *-og^u*, which is the majority plural for native class XI nouns. But in other cases, borrowings systematically select a minority plural, one that I have assumed is listed in the lexical entry of native nouns. So *n*-final borrowings not referring to male persons or roles systematically take the minor additive plural *-ab*, rather than the local default plural for class VI, replacive *b* (e.g. borrowed *tin* ~ *tinab* 'can, tin' (**tib*), *sipun* ~ *sipunab* 'spoon' (**sipub*); cf. native *aun* ~ *aub* 'sun, moon').

This observation does not affect the argument for the peripherality of *-ehas* among the Arapesh plurals, since what is crucial to that argument is simply the fact that pluralization in borrowings is sensitive to the noun's phonological form, which is true in any event. It does, however, constitute a problem given the specifications I have assumed for Arapesh lexical entries, since only the majority form-based plurals are supplied by a feature-filling rule. Representational revisions are conceivable. For instance, we might assume that it is the additive suffixes that are the basic morphological plurals, with the apparent replacives lexically marked to undergo a readjustment rule so that $Xt \rightarrow Xtog^u \rightarrow Xg^u$. While this has an advantage in the description of loanword behavior that will become apparent in a moment, it has problems as well. In particular, it requires lexical marking in the vast majority of nominal entries for this extra rule, since XI is the only class in which the additive alternant (*tog^u*) is also the majority plural. If we are willing to accept this degree of lexical marking, we might as well list the plural of every noun in its lexical entry, and do away with the notion of a local default plural completely. In that case *all* the plural rules would be redundancy rules. I will consider this possibility again later on.

I would like to approach the problem from a different angle at this point, however, in light of a generalization about which form of the plural

loanwords *do* tend to select. As we have seen, when there is a choice of alternate forms with which to realize the plural of a borrowed noun, the grammar strongly favors the long form, i.e., the additive suffix, over the replative. Since this generalization is at odds with the one that holds for native nouns, the operation of a separate principle is implied. It might be stated most straightforwardly as follows: ‘do not obscure morphological boundaries’. One could formulate such a principle as an Align constraint calling for the coincidence of morphological and phonological boundaries in an optimality-theoretic grammar (see e.g. McCarthy and Prince 1993a,c). However, since I am interested in the constraint’s substantive properties, I will appeal instead on the notion of ‘morphotactic transparency’ developed within the theory of natural morphology (Mayerthaler 1981, Wurzel 1989), according to which sound-meaning correspondences that are transparent are evaluated as relatively less marked and therefore relatively more natural, all else being equal. As Kilani-Schoch and Dressler explain it (1984:51, see also Dressler 1985), ‘the more transparent the base of the form is, the better perceptible it is and the better is the motivation of the derived word or word form. Morphotactic transparency is total if affixation does not change anything at the boundary between base and affix’.

I propose that the tendency to pluralize borrowed nouns with the additive rather than replative ending is a direct reflection of this preference for morphotactic transparency. Such a preference surely plays some role in the native morphology as well, but it can be particularly influential in the realization of loan words precisely because they are new members of the lexicon and so are not affected by the phonological wear of use over time (Bybee 1994). The tendency for loans to select the additive plural might be construed as another instance of ‘the emergence of the unmarked’ (McCarthy and Prince 1993b), the idea that general constraints may be elevated in status in marked contexts, in this case, in loan vocabulary where ease of recognition of the singular form is especially helpful.²¹

There is some independent justification for invoking a principle of morphotactic transparency for loanwords rather than revising our entries for native Arapesh nouns. One piece of evidence comes from the class IV borrowing *tommihawk*^u ‘tomahawk, trade axe’, which is pluralized as *tommihaguhijer*, i.e. with a relatively uncommon plural for native class IV

²¹This is similar to the reasoning Wegener 1995:249, 2002 uses to explain why we regularly find the *s*-plural on ‘marked’ German nouns, a class which includes proper names, clippings, foreign words with a marked final vowel, quoted words, and onomatopoeic words: ‘The *s*-plural protects these marked nouns from too substantial changes of their sound structure... because in contrast to the syllabic plural inflections, the *-s* does not cause a shift of the syllable boundary of the singular stem,... and because it excludes umlaut’.

nouns. The reason for this appears to be that *guhijer* is the one class IV plural that preserves the place and manner of the final consonant in the corresponding singular, making it the alternant of choice in terms of morphotactic transparency.²²

The other source of evidence is even more compelling. Native class X nouns, which end in *l* or *r* in the singular, select only one plural allomorph, namely replacive *guh*: *uruwhir* ~ *uruwhigu^h* ‘fish net’, *efur* ~ *efugu^h* ‘sago grub’.²³ Of the two liquid-final borrowings in Fortune’s data, one pluralizes as expected (*neil* ~ *neigu^h* ‘nail’). But the other surprisingly takes an additive ‘long-form’ plural, even though no such plural exists for native nouns: *botar* ~ *botaragu^h* ‘bottle’. Similarly, in Cemaun Arapesh the Tok Pisin borrowing *hauskuk* ‘cooking house’ receives the form-based class IV plural *meb*, but unlike any native *meb*-plural, it is not replacive: *hauskukmeb* (cf. the regular replacive pattern *ñagiruk* ~ *ñagirumeb* ‘type of bird’). Thus we see that the pressure to maintain morphotactic transparency of the base form can influence plural selection in a way that is demonstrably independent of the lexical specification of native nouns.

In any case, it should be clear that *-ehas* plays a qualitatively distinct role in Arapesh number marking. Unlike all the other plurals, *-ehas* is free of phonological restrictions on what it can attach to, which makes it the alternative chosen when no plural can be assigned on the normal basis of a noun’s singular form.²⁴ Moreover, the strong correlation between *-ehas* and class VIII syntax is not representative of the relation between plural realization and noun class in the language more generally. The regular pattern is rather for noun classes to be associated with phonologically specifiable categories of singular form. It is only when a noun’s form bears no predictable

²²The only other plural with this quality is the additive plural suffix *-omi*, which has a highly limited distribution: it occurs on only three native class IV nouns, the kinship terms for ‘mother’, ‘mother’s sister’, and ‘grandmother’. This same additive plural appears as a plural of a few class VII male kin terms as well. In any case, there are no instances of the *-omi* plural on native nouns that refer to non-humans, so the fact that it does not appear as the plural of *tommihawk^u* is not surprising.

²³There is a single class X noun, *mbul* ~ *mbulugu^h* ‘pig’, that takes an additive form of the plural. Although according to Fortune (1942:34) *mbul* is ‘the hardest worked member of the class,’ this word may itself be a borrowing, since it ‘is also isolated in a solitary use of the *mb* sound in the language’, and as Fortune notes, is ‘common to very many neighboring languages in its singular form’.

²⁴That is assuming that some plural must be selected, which seems to be the case for the material considered here. It is not necessarily true in all presently spoken varieties of Arapesh, however. In Cemaun there are a few nouns, both native and borrowed, which simply receive no plural, and which take not default agreement but rather agreement based on their phonological form (within the limits of the native agreement morphology). These are discussed in Chapter 5.

relation to its plural (which is the case for nouns pluralized with *-ehas*) that class VIII is called upon in its default capacity.

Where does this reanalysis leave the morphology-by-itself model, according to which phonological form, purely morphological realization class, and morphosyntactic noun class or gender are hierarchically ordered levels of grammatical structure? The model predicts that direct mappings between noun class and phonology should not be systematically exploited by a grammar, but not only are such mappings found, they even constitute the normal case in Arapesh. Furthermore, there is no revised ordering of the levels that would enable us to express the essential generalizations we wish to make: that for the large majority of Arapesh nouns class is rooted directly in noun-final phonological form, that this correlation is constrained to hold only for those cases in which the relation between phonological form and realization class is also predictable (this is what Aronoff means to capture by routing noun class assignment through plural realization), and that in at least one important case (nouns suffixed with *-ehas*) plural realization correlates with noun class directly. Any level we assigned to the intermediate position would always be required to admit exceptions. It would not necessarily constrain the mapping between levels, providing us with no good way to express the observation that noun class correlates with singular form, but takes simultaneous account of the plural.

The solution, of course, is to dispense with the notion of levels as the correct way of modeling the phenomenon, and directly acknowledge the distinct sorts of information that are involved in noun class assignment without placing architectural limits on the way they may correspond. This in no way compromises the key insight that the Arapesh plural realization and agreement classes are ‘autonomous’, not reducible to some other aspect of grammar and following principles of their own. Plural realization is just one among many sorts of morphologically-relevant information associated with the noun lexeme. Salient lexical properties in a number of dimensions—not only conceptual and semantic, not only morphological, but also phonological—can provide the basis for the classification of nouns.

The Lexical Representation of Arapesh Nouns

[M]y dear cannibals, the Buna of Marienberg... with their nine noun classes and object suffixes and other nice little accessories associated with them, have the most monstrous way of speaking known to me. If only these cannibals... would at least allow their nine classes to be arranged according to definite criteria. But no; apart from finding natural gender distinctions, at any rate, I have given up hope of it. I absolutely cannot grasp why, e.g., the fingernail goes into Class A, the finger goes into Class B, the hand goes into Class C, the lower arm goes into class D, the upper arm goes into class E, the entire arm goes into Class F, etc. Of course, this is a random example; in reality the 'arbitrarinesses'—which originally were surely not arbitrary—come if anything even more blatantly to the fore. I have gone to their 'class-brothers', the Bantu in Africa—but to no avail. I have tried to blame phonological influences, but these, too, are not responsible. I have appealed to the spirits, but so far even these have left me in the lurch. There is thus nothing left to do but learn each individual word according to its use.

—Fr. Franz Kirschbaum¹

3.0 Introduction

We may never know exactly what the exasperated missionary Father Franz Kirschbaum had in mind when he wrote the above description of the Torricelli language Buna (Laycock 1973), since if he ever wrote a grammar of the language it appears to have been destroyed along with most other mission records during World War II, and since as far as Don Laycock was

¹Kirschbaum 1921/22:1052-53 (translated from the German by the author).

able to tell from a visit to the Marienberg area in 1971, all traces of noun classification seem to have vanished from the language. But as Laycock 1975:775 notes, the system Kirschbaum describes sounds remarkably like Arapesh, assuming Kirschbaum was taking the elements predictive of noun class to be class-marking affixes analogous to the noun prefixes of Bantu. Under that assumption, there would be no basis for assigning nouns to classes in Arapesh either, and the class of each noun would have to be learned ‘according to its use’, just as Kirschbaum found necessary for Buna.

In this chapter, I explore in more detail the problem confronted by Father Kirschbaum, justifying the assumptions made about the representation of Arapesh nouns in the preceding and following chapters. In Section 3.1, I argue that the noun-final elements that Fortune noncommittally calls ‘noun terminals’ are morphologically inseparable from the singular noun stems. Fortune makes little effort to support this analytical decision, though there is reason to question it; after all, the terminals are systematically associated with syntactico-semantic values (singular and plural number, morphosyntactic class, and in certain cases male/female human semantics), and to that extent they would seem to conform to the traditional notion of morphemes as recurrent partials of sound and meaning. Bound up with this issue in interesting ways is the question of how plurals are formed in Arapesh, in particular, whether a plural suffix is added to the singular form or replaces its final class-determining segment(s).

The analysis of Arapesh plurals presents its own set of problems. On the one hand, the fact that plurals are imperfectly predictable for many nouns necessitates listing the plural alongside the singular form in their lexical entries. On the other hand, since the form of the plural is normally highly constrained, the operation of plural-assignment rules is also implied. The solution we assumed in Chapter 2 for these conflicting observations will be expanded in the discussion in Section 3.2 below. Again we will see that an adequate analysis of Arapesh nouns requires us to recognize the systematic presence of redundant information in lexical entries.

3.1 The Lexical Representation of Singulars

In this section I evaluate the morphological status of the ‘terminal’ elements on Arapesh singulars, concluding that they are morphologically inseparable parts of singular nouns. Historically they may derive from class suffixes, as pairings of unmarked singulars with affixed plurals are well attested in situations where noun classes are being lost (see, e.g., Demuth, Faraclas, and Marchese 1986; Marchese 1988). But the question remains whether there is any justification for treating them as noun suffixes synchronically, perhaps analogous to the ‘word-markers’ *-o*, *-a*, and $-\emptyset$ posited by Harris 1991 for

Spanish. Clearly to its merit, an analysis positing noun-final word-markers—effectively noun-final morphemes—would allow some grammatical status to be assigned to the linguistically significant yet otherwise unrecognized ‘noun terminals’ of Arapesh. Treating the elements that are predictive of class as isolable morphological units would also allow the formulation of class assignment rules that map straightforwardly between these units and morphosyntactic features, nicely expressing the ‘morphomic’ ideal put forth by Aronoff 1994.

The published sources on the various Arapesh dialects are remarkably schizophrenic in their treatment of the noun class terminals, and few explicit arguments are given to support either position (i.e., whether the terminals are separate morphemes or form part of the noun stems). Fortune 1942:9 states unequivocally that the Arapesh ‘noun class terminals are, as far as the usage of the noun is concerned, an integral part of the root form of the noun,’ grounding that claim in the observation that nouns have ‘no detachable root form which may take the varying terminals of the varying classes.’ He does not designate noun terminals separately in his texts, as he does most agreement markers. At the same time, however, Fortune regularly lists noun terminals in isolation from particular nouns for expository purposes. A similar pattern is followed in Matthews’ n.d. grammatical notes on the Ilahita variety of Mufian. Gerstner’s grammar of coastal Arapesh devotes a section to categories of word formation that are expressed by suffixes (1963:4); included in it are noun class agreement marking and the formation of derived categories such as diminutives (*nogoteb* ‘knife’ ~ *nogoteb|igun* ‘little knife’), but not the formation of basic nouns. Gerstner recognizes both suffixation and replacement as ways of forming noun plurals, but even for plurals, noun-internal boundaries are never marked with the vertical line he uses to separate word-internal elements in his other examples.

The work of Robert Conrad and colleagues, by contrast, though explicitly modeled on Fortune’s analysis of Arapesh (see Conrad and Wogiga 1991:1, 9), treats the gender-predicting noun-final elements of Arapesh as obligatory suffixes. Thus, for example, in Bukiyip ‘[n]ouns are defined as a word class occurring with one of a closed set of 18 suffixes, the majority of which are differentiated for singular and plural’ (Conrad and Wogiga 1991:8). Nouns are assumed to have the (tagmemic) structure [+Nucleus +Number], as in the following example from Conrad and Wogiga (1991:9), where it is stated that ‘[t]he two different suffixes *-b* ‘class I singular’ and *-bis* ‘class I plural’ are the characteristic number suffixes which occur with class I nouns’:

- (1) *bi-b* *bi-bis*
 betel nut-I.SG betel nut-I.PL

Nevertheless, the marking of noun-internal structure is not continued throughout the grammar; it is dropped immediately after the short section that deals directly with noun morphology. A similarly reduced commitment to a root-suffix boundary is found in the Mufian examples in Alungum, Conrad, and Lukas 1978, in which morphemes are generally scrupulously glossed. In a 1986 description of his native dialect Abu', Otto Nekitel follows a similar practice: nouns are characterized by singular and plural 'suffixes' in the noun class charts, but the noun root/suffix distinction is not maintained consistently. For example, no noun suffix is implied by Nekitel's claim that 'a verb is influenced to agree with the noun... by a general rule that reproduce[s] either the ultimate or the penultimate consonant of the noun' (1986:185); note that Nekitel does not treat agreement as reproducing a noun suffix. On the other hand, a root-suffix configuration is clearly presupposed in the pluralization rule Nekitel offers for class VI (cognate with Fortune's class X): 'the [plural] rule causes the singular suffix /l/ to go to zero in the environment between a vowel and the plural suffix /kuh/' (1986:194). In sum, the primary sources offer an inconsistent picture of noun-internal morphological constituency.

There is only one extended argument for treating singular nouns as morphologically primitive rather than derived by obligatory suffixation of a number morpheme, the one offered by Aronoff 1992, 1994. Aronoff 1992:31 observes that if we set off the Arapesh singular-final terminals as suffixes, then for the roots that remain there is 'no phonological or other property... that is predictive of morphological [plural] class or of anything else'. In other words, by extracting out suffixes we introduce the new problem of determining how these suffixes are assigned to roots (Blevins 2006 offers a more general exposition of this problem). A similar issue arises in analyzing the theme vowels of Latin and Spanish verbs. The problem is particularly acute for Arapesh, since the classes are so numerous, and since semantic considerations are so rarely relevant to class assignment. This is just the problem that appears to have beleaguered Father Kirschbaum.

Aronoff also points out that the representations assumed for Arapesh singulars have consequences for the analysis of noun plurals.² For nouns

²Clearly I agree with Aronoff that plural forms are relevant to the analysis of singulars. But given that the claim we are trying to make is that grammatical properties of Arapesh nouns (noun class, plural realization) are highly predictable on the basis of phonological form, it is our obligation to provide independent justification for the phonological forms we posit. The reasoning Aronoff 1994:109 provides is that extracting singular suffixes creates a fault line

that have a replacive plural—where the singular-final terminal sound is replaced with a new terminal in the plural—the segmentation of singulars into root-plus-suffix proceeds smoothly. Examples of replacive plurals in Ilahita Arapesh from Matthews n.d. are presented in (2) below.

(2)	<u>Singular</u>	<u>Plural</u>	
<i>m ~ p</i>	<i>naem</i>	<i>naep</i>	‘eye’
	<i>ambem</i>	<i>ambep</i>	‘breadfruit seed’
	<i>bandufum</i>	<i>bandufip</i>	‘white shell’
<i>p ~ x</i>	<i>aolap</i>	<i>aolax</i>	‘bush house’
	<i>okwatip</i>	<i>okwatix</i>	‘banana leaf’
	<i>sambap</i>	<i>sambax</i>	‘hand’
<i>t ~ ngw</i>	<i>nombat</i>	<i>nombangw</i>	‘dog’
	<i>dalot</i>	<i>dalongw</i>	‘cheek’
	<i>gweit</i>	<i>gweingw</i>	‘small axe’

But as discussed in Chapter 2, in many cases the plural is formed by the simple addition of a suffix, with the singular-final segment retained in the plural. Examples of this pattern (again from Matthews) are presented in (3).

(3)	<u>Singular</u>	<u>Plural</u>	
<i>ga ~ gax</i>	<i>balanga</i>	<i>balangax</i>	‘head’
	<i>ontanga</i>	<i>ontangax</i>	‘cloud’
	<i>sobunga</i>	<i>sobungax</i>	‘kapok tree’
<i>mb ~ mbex</i>	<i>kofemb</i>	<i>kofembex</i>	‘egg’
	<i>bemb</i>	<i>bembex</i>	‘betel nut’
	<i>nomamb</i>	<i>nomambex</i>	‘breast’
<i>na ~ nab</i>	<i>ələlina</i>	<i>ələlinab</i>	‘rainbow’
	<i>aunguna</i>	<i>aungunab</i>	‘egret’
	<i>geina</i>	<i>geinab</i>	‘rifle’

The isolated Bukiyip example in (1) above, which purports to demonstrate the structure [+Nucleus +Number] is thus misleading; it implies that the suffix-initial *b* in both forms is coincidental when in fact it is a thoroughgoing pattern across many noun classes. Indeed, in Arapesh the pro-

through the predictability relationship between phonological form and class that corresponds exactly to the replacive/non-replacive plurals: we would only be able to ‘posit phonological predictors of [realization] class... [where] the segmentation procedure results in a zero singular suffix’. While true, this statement is not a sufficient argument against singular suffixes, since it is with this same predictability that we would like to conclude.

posed ‘singular suffix’ is preserved not only in the plurals of *b*-final nouns, but also with subclasses of both *g*- and *t*-final nouns, as in the (IIIb) and (XIb) forms from Rohwim in (4). As discussed in Chapter 2, these alternative ways of forming the plural are purely realizational, and have no bearing on a noun’s morphosyntactic classification.

(4)	<u>Singular</u>	<u>Plural</u>	
IIIa	<i>deiwig</i>	<i>deiwas</i>	‘latrine place’
	<i>barawag</i>	<i>barawas</i>	‘spear’
IIIb	<i>abeg</i>	<i>abegas</i>	‘betel nut quid’
	<i>bog</i>	<i>bogas</i>	‘chopstick for winding sago’
XIa	<i>nybat</i>	<i>nybag^u</i>	‘dog’
	<i>unut</i>	<i>unug^u</i>	‘bundle’
XIb	<i>wit</i>	<i>witog^u</i>	‘door’
	<i>fokwet</i>	<i>fokwetog^u</i>	‘menstrual house’

It should be emphasized that while additive plurals may be somewhat less common than replatives in Arapesh, they are by no means rare. The additive pattern is also the one most often found with loan words as described in Chapter 2. If the singular-final elements were supplied by morphological rules triggered by a plural feature on roots, then we would be forced to either account for the application of these rules in the plural forms of the many nouns like those in (3) and the IIIb and XIb subclasses of (4) above, including their productivity in loan words, or else accept that a considerable subset of plural formatives just so happen to contain the corresponding singular morphemes at their left edge.

By contrast, assuming the noun singulars to be morphologically primitive has relatively few negative repercussions.³ In fact, it even has a positive result independent of the issue of predicting class, which can be explained by analogy to Corbett’s (1991:20-24, 64-65) discussion of a semantic gender rule in Algonquian. The two Algonquian genders, conventionally characterized as ‘animate’ and ‘inanimate’, have two strong associations: one, there is a strong tendency for nouns whose referents are considered powerful to be treated as grammatically animate, and two, animate gender correlates virtually exceptionlessly with singular/plural suffixes that appear obligatorily on nouns. The pattern is represented in (5) below with examples from Fox (Dahlstrom 1995).

³The main consequence is the need for replative pluralization rules. These are discussed below.

(5)		<u>Singular</u>	<u>Plural</u>	
	<u>(Proximate) Animate</u>	<i>mahkw-a</i>	<i>mahkw-aki</i>	‘man’
		<i>neniw-a</i>	<i>neniw-aki</i>	‘bear’
	<u>Inanimate</u>	<i>mehtekw-i</i>	<i>mehtekw-ani</i>	‘wood’, ‘tree’, ‘stick’
		<i>menes-i</i>	<i>menes-ani</i>	‘island’

Although semantics alone is not a sufficient predictor of gender, whereas the morphology is, Corbett (1991:65) argues for a semantic determination of gender wherever possible, because

lexical entries must include information as to meaning and phonological form; claims that they include other elements require justification.... If we were to treat [Algonquian] as having morphological assignment, then the declensional type of each noun would have to be indicated in the lexicon.... The assignment rule would apply without exception, but the morphological feature would be redundant for almost every noun.

In other words, we should take into account what has to be learned about an item for independent reasons. Since there are so few general rules associating Arapesh noun terminals with stems, the terminals (or a diacritic) would have to be memorized for all but the human nouns. But under an analysis in which singulars are treated as morphologically simple, what has to be learned is simply the noun’s phonological representation, a representation that is typically only one segment longer (and at most only three) than the noun ‘root’ on the interpretation that the class-predicting noun-final segments are suffixes.

Assuming that singular forms are indivisible units also appears correct in light of Fortune’s comment cited above that noun ‘roots’ are not reused with other ‘suffixes’ within the system. The few noun roots that are found to have consistent meaning with alternate terminals are kinship nouns in which the terminals are associated with a semantic value (male or female human) in addition to marking noun class, as in the examples in (6).

(6)	<u>Male (VII)</u>	<u>Female (IV)</u>	
	<i>arapeñin ~ arapeñin</i>	<i>arapeñik^u ~ arapeñijer</i>	‘friend’
	<i>megan ~ meganomwi</i>	<i>meganukik^u ~ meganubiheu</i>	‘same-sex in-law’
	<i>afuken ~ afukenim</i>	<i>afukek^u ~ afukeu</i>	‘same-sex elder sibling’
	<i>sigaben ~ sigabehem</i>	<i>sigabehik^u ~ sigabeheu</i>	‘ghost’

In these cases, it seems reasonable enough to posit a root-suffix structure along the lines of Aronoff's 1994:112 treatment of *k^u*-final singulars referring to human females. In the case of transparently derived diminutives, suffixes are not added to a root-like constituent but rather to the singular form in its entirety, regardless of the mode used for forming the plural, as shown in (7).^{4,5}

(7)	<u>Singular ~ Plural</u>	<u>Diminutive</u>	
Abu'	<i>utam ~ utaba</i> <i>dubaun ~ dubaub</i> <i>aleman ~ alemam</i>	<i>utamikil</i> <i>dubaunikin</i> <i>alemanikin</i>	'stone' 'crayfish' 'man'
But Arapesh	<i>nogoteb ~ nogoges</i> <i>ipab ~ ipas</i>	<i>nogotebigun</i> <i>ipabigun</i>	'knife' 'banana'
Bukiyip	<i>bul ~ bulguh</i>	<i>buligin</i>	'pig'

In Cemaun, the noun-final elements associated with classes are also put to use as suffixes individuating non-count nouns, and in these cases, too, there may be some semantic association with the suffixed class marker. So for example *kwiñ* 'stirred and spun sago', a mass of jelly-like substance, is individuated into *kwiñ-i-b* 'balls of sago jelly' (cf. *ñumeb* 'breasts'). Similarly, *cukwehəbər* 'type of edible mushroom' refers to the mass of mushrooms that grow on piles of chipped sago pith left to rot in the jungle after the starch has been washed out. To refer to one piece, *cukwehəbər-i-p* 'mushroom-POSS-XI.SG' is used, with the final class marker referencing the implied possessed noun *cup* 'leaf'.

But to characterize the entire set of singular Arapesh nouns on the basis of these relatively peripheral semantically compositional cases would con-

⁴Contrast the Arapesh pattern with Swahili:

<u>Class</u>	<u>Noun</u>	<u>Diminutive</u>	
<i>m-/wa</i>	<i>m-toto</i>	<i>ki-toto</i>	'child'
<i>N-/N-</i>	<i>m-buzi</i>	<i>ki-buzi</i>	'goat'
<i>m-/mi-</i>	<i>m-lima</i>	<i>ki-lima</i>	'mountain'

Alternation between classes may be meaningfully exploited in noun classification systems as a method of word formation (see Denny 1976, Mufwene 1980, Dixon 1982, Serzisko 1982, Corbett 1991, Contini-Morava and Kilarski 2011). In such cases the class marker is typically replaced rather than augmented.

⁵The But data is from Gerstner 1963. In Cemaun, there is no one suffix that is consistently used to mark diminutivity. While the locative suffix *-gin* is commonly used for this purpose (*bur-i-gin* 'pig-POSS-LOC'), many nouns suffix class markers on a lexical basis: *bur-i-k^w* 'pig-POSS-IV.SG', *arəmir-i-p* 'bird-POSS-XI.SG', *nimbat-i-h^w* 'dog-POSS-XII.SG'. Each of these suffixes triggers phonological agreement. Agreement follows the suffix and not the base.

flate the derivational and inflectional functions of the terminals and misrepresent the system. In by far the majority of nouns, class ‘suffixes’ have no consistent meaning, and attach to ‘roots’ which have no consistent meaning either. Indeed, there is no more independent justification for a root-stem structure in words like Arapesh *ulypat* ‘house’ (i.e., *ulypa + t) than there is in the word *house* itself in English (i.e., *hau + s). For a demonstration, consider the following root + suffix combinations in Cemaun, which were hypothesized for nine ‘roots’ in each of the twelve classes that are semantically available (i.e., nouns were hypothesized for all classes but the exclusively masculine human class VII):

(8)	<u>Root</u>	<u>Singular</u>	
	bu-	<i>buk</i>	‘riverbank’
		<i>buh^w</i>	‘tangled treetops’
		<i>bur</i>	‘pig’
		<i>bum</i>	‘sprout’
		<i>but</i>	‘type of tree’
		<i>*bub, *bun, *buñ, *bup, *bug, *bubir, *buk^w</i>	
	aha-	<i>ahat</i>	‘muddied water’
		<i>ahah</i>	‘finger’
		<i>aham</i>	‘tongue’
		<i>ahabir</i>	‘taboo period after a death’
		<i>ahar</i>	‘type of tree’
		<i>*ahap, *ahab, *ahan, *ahag, *agañ, *ahah^w, *ahak^w</i>	
	ya-	<i>yah</i>	‘road’
		<i>yañ</i>	‘tiny frog’
		<i>yap</i>	‘type of tree’
		<i>*yab, *yan, *yam, *yak^w, *yar, *yat, *yag, *yah^w, *yabir</i>	
	api-	<i>apik</i>	‘type of yam’
		<i>apib</i>	‘knot’, ‘loop’
		<i>*apig, *apir, *apin, *apim, *apit, *apip, *apih, *apuk^w, *apuh^w, apiñ</i>	
	ma-	<i>mah</i>	‘voice’
		<i>mag</i>	‘type of tree’
		<i>map</i>	‘narrow end of a tuber’
		<i>*mañ, *mat, *man, *mab, *mam, *mar, *mah^w, *mak^w, *mabir</i>	

wa-	<i>wah</i>	‘sunlight’
	<i>wab</i>	‘night’
	<i>wabir</i>	‘village’
	<i>wap</i>	‘nest’
	<i>*wan, *wañ, *wam, *wat, *wak^w, *wah^w, *wag, *war</i>	
na-	<i>nat</i>	‘small game trap’
	<i>nabir</i>	‘fence’
	<i>nah^w</i>	‘tooth’
	<i>*nap, *nab, *nar, *nan, *nak^w, *nah, *nag, *nañ, *nam</i>	
abori-	<i>aborir</i>	‘shadow’
	<i>aborim</i>	‘tree used to make children’s bows’
	<i>*aborig, *aborin, *aboriñ, *aborib, *aborip, *aborih^w, *aborit, *aborik^w, *aboribir, *aborih</i>	
su-	<i>sur</i>	‘bush camp’
	<i>sut</i>	‘house post’
	<i>suk^w</i>	‘liver’
	<i>*sup, *suh^w, *suh, *sum, *subir, *sub, *sun, *suñ, *sug</i>	

The ‘roots’ *wa-* and *na-* have at least some potential unity of meaning; the *wa-* words have to do with light or home, and two of the three *na-* words have to do with things that stand in rows. But the associations are tenuous at best. Indeed, these ‘roots’ hardly differ from what we would get from a comparable exercise in English in degree of semantic coherence and distribution. Compare the four English noun ‘families’ hypothesized in (9).

(9) Root

/bi-/	<i>beet/beat, beak, bead, bean, beam, beep, beach, beef, beer, beast, *biece, *beague, *beeb, *biend, *beaje, *beeve, *beeze</i>
/kə-/	<i>cut, cuss, cup, cub, cud, come, cuff, cusp, cunt, *cuck, *kug, *kunn, *kung, *kutch, *kudge, *kuv, *kump, *kunk, *kunch, *cuzz, *cust, *kush</i>
/grei-/	<i>grate, grace, grape, grade, grain, grave, graze, *grake, *grame, *grage, *grafe, *grraith</i>
/ræ-/	<i>rash, rat, wrath, rack, rag, wrasse, wrap, ram, rant, ramp, rank, ranch, rasp, *rab, *rad, *ran, *rang, *ratch, *radge, *raff, *rand, *rax, *razz, *rast</i>

If we propose that Arapesh has a root *wa-* having to do with light and home, there is nothing to prevent us from also positing an English root *kə-* having to do with edges (*cuff*, *cusps*) and curse words (*cuss*, *cunt*).

Nevertheless, certain terminals remain problematic. There is a class of nouns, present in cognate form across the family, that end in the sequence *bVr* in the singular and that take the metathetic plural *rVb* (10).

(10)	<u>Singular</u>	<u>Plural</u>	
Rohwim	<i>wabør</i>	<i>waryb</i>	‘village’
	<i>ñemabør</i>	<i>ñemarøb</i>	‘string’
	<i>ipokwibør</i>	<i>ipokwibryb</i>	‘side over kidneys’
Abu’	<i>wabul</i>	<i>walub</i>	‘village’
	<i>numabul</i>	<i>numalub</i>	‘twisted net bag string’
	<i>asabul</i>	<i>asalub</i>	‘singing’, ‘traditional dance’
Mufian	<i>numambel</i>	<i>numalemb</i>	‘thread’
	<i>endambel</i>	<i>endalemb</i>	‘fence’
	<i>nembel</i>	<i>nelelemb</i>	‘stomach’
But Arapesh	<i>wabul</i>	<i>walub</i>	‘village’
	<i>nomabul</i>	<i>nomalub</i>	‘string’, ‘fishing line’
	<i>manyebul</i>	<i>mat’ulub</i>	‘root’
Bukiyip	<i>wabəl</i>	<i>walib</i>	‘village’
	<i>nabəl</i>	<i>nalib</i>	‘fence’

Denying the noun terminals morphemic status is one thing when the terminals are simply segments. But it is harder to justify for a recurrent substring that can even be removed, restructured, and replaced in the service of morphological marking. Yet to the extent that these polysegmental terminals fail to attach to an identifiable root and have no specifiable semantics, they are no better candidates for a suffixal analysis than are the single-segment terminals.

A hybrid solution would be to recognize something like the ‘dotted-line’ lexical entries adopted by Platt 1981 for Old Provençal verbs or by Janda 1987:466-467 for Latin nouns. Janda’s analysis of Latin combines noun stems with their thematic vowels, which are in turn associated more or less directly with nominal declension class. If the nouns have complex lexical entries as in (11), it is not necessary to assume that the correct theme vowel and thereby the correct set of inflectional suffixes are channeled onto them by some kind of diacritic marking (i.e., using diacritic features like [+declension 1] or [+class II]).

(11)	Latin:	/port+a/	<i>porta</i>	‘gate’ (declension 1)
		/faci+e:/	<i>facie:s</i>	‘face’, ‘appearance’, ‘shape’ (declension 4)
	Arapesh:	/ñema+bør/	<i>ñemabør</i>	‘string’ (class II)
		/barawa+g/	<i>barawag</i>	‘spear’ (class III)

In Arapesh, the morphologically active noun-final element would likewise be attributed to a distinct part of the representation. Under this view, the fact that these elements are so closely associated with class is sufficient to confer morphemic status upon them, even though they are lexically dependent upon the singular stems. It is not clear to me, however, whether this interpretation can be differentiated empirically from the non-composite one I have been assuming all along.⁶

The ‘automorphology’ algorithm of Goldsmith 2001 would attribute morphemehood to the Arapesh singular noun terminals because they pattern into paradigms with their corresponding plural forms, constituting ‘signatures’, such as $\{-t, -g^w\}$ or $\{-r, -guh\}$, that recur on multiple ‘stems’. Reference to a noun’s meaning—whether its lexical semantics or its morphologically contrastive ‘information content’ (Carstairs-McCarthy 1994) would not be relevant. Under such a conception of morphemehood, the fact that most Arapesh terminals are semantically empty and associated with stems on a purely lexical basis would not prevent them from being identified as morphemes. However, the result would be morphemes of a rather different nature than the ones we have been testing for here. To the extent that the analysis would establish a large and thoroughgoing system of inflectional cranberry morphs (which is what we would end up with), it calls into question any definition of morpheme that requires reference to meaning.

As a final piece of evidence that Arapesh singular nouns are morphological primitives, we can return to the argument made in Chapter 1 for Abu’ (to be developed further in Chapter 5) that the number of noun classes approaches the number of phonotactically permissible word-final consonants. This generalization cannot be pushed to its absolute limits. The

⁶Zellig Harris’s interpretation of Swahili noun class prefixes as ‘discontinuous morphemes’ is based on this same idea: ‘It is true that we discuss the phonemes independently of the morphemes, and similarly we should discuss the repeated initials [i.e., the class prefixes] independently of the individual nouns. But we should discuss them merely as a feature of the phonemic composition of a class of Swahili morphemes, namely the fact that these morphemes are discontinuous and that the discontinuity consists in the repeating of the initial (or of a substitute for it) in stated positions throughout a certain [syntactic] domain. All this does not require us to set up the repeated initial as a separate morpheme’ (1945:127).

‘whispered terminals’ or colored aspiration Fortune transcribes partition final consonants into more categories than the morphology does, and the phonological and morphological classifications do not perfectly correspond.⁷ But to the extent that it does hold—and for the most part it does—we would be forced to posit a singular suffix that corresponds to nearly every word-final consonant in the language. If, on the other hand, we assume that the noun terminals are simply a part of the nouns’ phonological representations, the close correlation between phonotactically permissible noun-final sounds and noun class categories falls out as a natural consequence.

3.2 The Lexical Representation of Plurals

In the previous chapter we saw that the Arapesh noun plurals are canonically associated with phonologically defined classes of singular noun form. But although this statement is generally true, there are many irregularities in plural realization that complicate the relationship. Because formally and morphosyntactically similar plurals need to be individually listed in some cases, yet can be supplied by inflectional realization rules in others, ‘the same’ plural ‘morpheme’ is often distributed across both rules and representations. As we have shown, class assignment makes crucial reference to these elements irrespective of their listed or rule-derived status, and so some mechanism is needed for unifying the clearly related yet irreducibly distinct plural alternants which coexist for many classes of Arapesh nouns.

Here I will illustrate several patterns of Arapesh plural formation, beginning with examples from Abu’ which are relatively regular in the overall Arapesh context. We will then consider other plural patterns in order of increasing irregularity, including a formally identifiable class of Arapesh plurals in which most if not all items must arguably be listed, despite being partially predictable.

Consider the Abu’ nouns in (12), which are treated as a class by Nekitel 1986 (their Arapesh cognate class is VI).

(12)	<u>Singular</u>	<u>Plural</u>	
	<i>ehen</i>	<i>eheb</i>	‘cold river fish’
	<i>buburan</i>	<i>buburab</i>	‘coconut shell ladle’
	<i>dubaun</i>	<i>dubaub</i>	‘lobster’, ‘crayfish’
	<i>kedin</i>	<i>kedib</i>	‘softwood tree’

⁷For example, labialized *k^u* is referenced by the noun class system, whereas non-labialized *k* is treated as an exception.

These nouns conform to the simple plural formation rule in (13). The rule is replacive; that is, the plural marker replaces the singular noun terminal.

- (13) +PLURAL
 $Xn \rightarrow Xb^8$

The analysis for this class is complicated, though not hopelessly, by the second sub-pattern shown in (14).

(14)	<u>Singular</u>	<u>Plural</u>	
	<i>ain</i>	<i>ainab</i>	'iron' (TP)
	<i>pin</i>	<i>pinab</i>	'pin' (TP)
	<i>baten</i>	<i>batenab</i>	'button' (TP)
	<i>keina</i>	<i>keinab</i>	'bow' (Ulau-Suain) ⁹
	<i>titiakamuna</i>	<i>titiakamunab</i>	'millipede'

Only *titiakamuna* appears to be a native noun. This group differs only slightly from the nouns in (12); here the plural is suffixed directly to the singular form, rather than replacing the final *n*. In Chapter 2, we decided to treat Arapesh replacive plurals as supplied by rule when they were the majority form, and account for the productive application of additive plurals in loan words by invoking a supplementary principle of morphotactic transparency, which prevents replacement in borrowed vocabulary because to do so would obscure a morphological boundary. Presumably the preservation of singular-final *n* in the examples in (14) is likewise due to this principle. Assuming with Nekitel 1992 that *a* is epenthetic, then the assignment of plurals for nouns of this class will be determined as in (15).

- (15) +PLURAL
 $pin \rightarrow *pib$ (morphotactic transparency)
 $\rightarrow *pinb$ (illicit CC# cluster)
 $\rightarrow pinab$ (repair via epenthesis)

If this is basically correct, then a unitary plural rule can be maintained for the class of Abu' $n \sim b$ nouns.

⁸There is only one slightly irregular form in this set, *dubaren* ~ *dubarub* 'hornbill'.

⁹Ulau-Suain is an Austronesian language spoken to the west of the Arapesh region.

The next set of morphosyntactically equivalent plurals, represented in (16), is cognate with Arapesh class IX. The plurals conform to the replacive rule in (17).

(16)	<u>Singular</u>	<u>Plural</u>	
	<i>asaf</i>	<i>asas</i>	‘pubic covering’
	<i>du’u’naf</i>	<i>du’u’nas</i>	‘tree with edible leaves’
	<i>idaf</i>	<i>idas</i>	‘fencing timber’
	<i>ihiaf</i>	<i>ihias</i>	‘sliced taro/yam’
	<i>nabulaf</i>	<i>nabulas</i>	‘tree used for making drums’
	<i>suaf</i>	<i>suas</i>	‘stone for sharpening knives’

(17)	+PLURAL	
	<i>Xf</i> →	<i>Xs</i>

Some plurals of this class have fronted final vowels (18).

(18)	<u>Singular</u>	<u>Plural</u>	
	<i>aluf</i>	<i>alis</i>	‘body’
	<i>rabuf</i>	<i>rabis</i>	‘liver’
	<i>o’o’suf</i>	<i>o’o’sis</i>	‘bottom’

Since the variation in plural realization appears to be regularly conditioned, we might posit a pair of rules, *uf* → *is* and *f* → *s*; alternatively, a readjustment rule might be called upon to bring about the vowel change in *uf*-final nouns. At the very least, the plural will have to be listed for the similarly classed word for ‘liver’, *a’uf* ~ *a’uis*, where the *u* is unexpectedly preserved in the plural.

In both of the above cases, a unified plural rule can be maintained for nouns belonging to the same morphosyntactic class. Equally representative, however, is the pattern found in (19), which corresponds to Arapesh class XII.

(19)	<u>Singular</u>	<u>Plural</u>	
	<i>awah</i>	<i>awalih</i>	‘song’
	<i>bur’ah</i>	<i>bur’alih</i>	‘leg’
	<i>lahuh</i>	<i>lahulih</i>	‘sago’
	<i>lah</i>	<i>lalih</i>	‘gouging utensil’
	<i>maduh</i>	<i>madulih</i>	‘vine’, ‘rope’

Accounting for the plurals of these nouns requires a familiar replacive rule (20).

- (20) +PLURAL
Xh → *Xlih*

In addition, however, this class includes many nouns with diverse irregularities in plural formation (21).

- | | | | |
|------|--------------------------|------------------------------|--------------------------------|
| (21) | <u>Singular</u> | <u>Plural</u> | |
| | <i>ihia^{uh}</i> | <i>ihia^{urilih}</i> | ‘butterfly’ (V change) |
| | <i>halihif</i> | <i>helih</i> | ‘feather’ (V change, deletion) |
| | <i>nukofuh</i> | <i>nukosilih</i> | ‘navel’ (C change, V change) |

Some of the stem-internal alternations in (21) correspond to number alternations found noun-finally elsewhere in the language; the *f*~*s* alternation, for example, is the canonical pattern described in (17) above (this is the phenomenon of ‘intervocalic change’ to be discussed further in Chapter 4). Others, such as *halihif*~*helih* ‘feather’, are more idiosyncratic; here it is the singular that is formally odd. Since there is no way to predict which nouns will be subject to stem-internal alternations, nor which alternations they will be subject to, we can only assume that they must be learned on a case-by-case basis.

As we turn from Abu’ to Rohwim, we find a still greater degree of unpredictability in plural formation. The examples in (22) represent the two main patterns found in Arapesh class XII plurals.

- | | | | |
|------|--------------------------|-------------------|--------------------------------------|
| (22) | <u>Singular</u> | <u>Plural</u> | |
| a. | <i>age^{uh}</i> | <i>agegwiruh</i> | ‘cleared space in center of village’ |
| | <i>nuba^{uh}</i> | <i>nubagwiruh</i> | ‘house for keeping sacra’ |
| | <i>wano^{uh}</i> | <i>wanogwiruh</i> | ‘ambush’ |
| b. | <i>pawe^{uh}</i> | <i>paweruh</i> | ‘chip’ |
| | <i>sora^{uh}</i> | <i>soraruh</i> | ‘ground sliding after rains’ |
| | <i>aro^{uh}</i> | <i>aroruh</i> | ‘slit gong drum stick’ |

The (22b) pattern is the more common of the two. It can be expressed by the general rule in (23).

- (23) +PLURAL
X^{uh} → *Xruh*

The appearance of the extra sequence *gwi* in the (22a) plurals is unpredictable, however. While it would be possible to posit a readjustment rule infixing this element into certain class XII nouns, there is no basis for determining which words it applies to, other than enumerating a list.

An alternative would be to posit two stems for each of these nouns, an unmarked stem and a *gwi*-final stem subcategorized for use with plural inflection. Perlmutter 1988 uses this approach to reanalyze the apparent presence of plural inflection inside the diminutive in Yiddish ablauting nouns in terms of stem suppletion. For Perlmutter, the Yiddish noun *ceplex* 'braid.DIM.PL', for example, is based not on the unmarked stem *cop*, but rather on a distinct 'ablaut' stem *cep*, which happens to serve as the plural form as well. To justify the Yiddish stem as an independent morphological unit, Perlmutter observes that for nouns that form their plurals by suffixation to an ablaut stem, such as *bux* 'book' ~ *bixer* 'books' and *boym* 'tree' ~ *beymer* 'trees', the diminutive is built directly on the ablaut stem, without the additional suffix that is present in the plural: *bixl*, *bixelex* (**bixerlex*) 'books.DIM'; *beymele*, *beymelex* (**beymerlex*) 'trees.DIM'. Stem form may be dissociated from number marking because some diminutives are formed from ablaut stems even when their plurals are built regularly; e.g., *hoz* 'hare' ~ *hezl*, *hezele* 'hare.DIM' ~ *hozn* (**hez*, **hezer*) 'hares'.

But in Arapesh the stem alternations would be relevant to nothing else besides the formation of plurals. There are few derivational markings on nouns that could provide evidence for the irregular stem as an independent morphological unit as in Yiddish. And for those that exist, such as the Abu' diminutive plurals in (24a,b), we find suffixation to either the singular or the plural form, but not to some other stem.¹⁰

(24)		<u>Noun</u>	<u>Diminutive</u>	
a.	SG	<i>nikam</i>	<i>nikamikil</i>	'taro'
	PL	<i>nikef</i>	<i>nikefikuh</i>	
	SG	<i>utam</i>	<i>utamikil</i>	'stone'
	PL	<i>utab</i>	<i>utabakuh</i>	
b.	SG	<i>dubaren</i>	<i>dubarinikin</i>	'hornbill'
	PL	<i>dubarub</i>	<i>dubarinikuh</i>	
	SG	<i>du'it</i>	<i>du'itikin</i>	'hill', 'mountain'
	PL	<i>disuk</i>	<i>du'itikuh</i>	

¹⁰Two irregular diminutives are also listed by Nekitel 1986: *raitawas* SG ~ *raitawasikuh* PL 'small portion of rice' (cf. TP *rais* 'rice'), and *lehitawas* SG ~ *lehitawasukuh* PL 'small portion of sago' (cf. *lehin* 'sago'). Here the base of diminutive formation is neither the stem nor the word.

The peculiarity of the class XII *gwiruh* plurals thus seems to be tied to nothing but the plural forms themselves.

Arapesh class XII looks more irregular still when we consider examples such as those in (25).

(25)	<u>Singular</u>	<u>Plural</u>	
	<i>aro^{uh}</i>	<i>araruh</i>	‘sugar cane’
	<i>arati^{uh}</i>	<i>arairuh</i>	‘man’s hair band’
	<i>raho^{uh}</i>	<i>rahororuh</i>	‘sago palm’
	<i>arukote^{uh}</i>	<i>aruweruh</i>	‘tongs’
	<i>awharuh</i>	<i>awhariruh</i>	‘fruit of breadfruit tree’

Here once again, although the alternation *kot* ~ *w* in ‘tongs’ resembles the noun-final *k^u* ~ *u* plural alternation of a large set of class IV nouns, it is unexpected word-internally, and the resemblance is only partial, since the noun shows an absence of *t* in the plural. Nor can any of these alternations be motivated phonologically. For example, the reduplication of stem-internal *r* in the plural form of *raho^{uh}* ~ *rahororuh* ‘sago palm’ cannot be motivated by a word-length requirement, given that *na^{uh}* ‘tooth’ pluralizes simply as *naruh*. The noun *aro^{uh}* ~ *araruh* ‘sugar cane’ has a homophonous minimal partner without the vowel change, *aro^{uh}* ~ *aroruh* ‘drum stick’ (see (22b)), so no phonological explanation will account for these either. These irregular plural forms have to be lexically specified. This is similar to the full suppletion analysis assumed by Perlmutter 1988:90 for Yiddish *-im* plurals, which he concludes must be listed due to the ‘high degree of irregularity in their singular-plural alternations.’ The formation of class XII plurals in Arapesh thus appears to involve one general rule, (23), and a handful of listed plurals that conform to that rule partially, but not entirely.

The central set of Arapesh class assignment rules, which we will refer to as morphological assignment rules, make simultaneous reference to both the singular and plural forms of nouns. Since plural formation rules should contain precisely the information that is required for this purpose, I have assumed that it is these rules that are critically accessed when morphosyntactic class is determined by morphological criteria. The material surveyed here makes it clear that a considerable number of morphosyntactically regular Arapesh plurals are listed rather than derived by the productive application of plural-formation rules, even when they conform to the expected plural-formation patterns to some extent. As a consequence, the rules referred to in class assignment must be redundancy rules of the sort illustrated in (26) below. These rules must be sufficiently general so as to allow listed plurals to be associated with their regular counterparts.

(26)	Abu'	$Xh \leftrightarrow Xlih_{[+PL]}$
	Arapesh	$Xuh \leftrightarrow Xruh_{[+PL]}$

The singular-plural alternations that are referred to by the regular morphological class assignment rules of Arapesh must therefore be interpreted as a set of lexical redundancy rules, rather than conventional feature-filling inflectional rules. Lexical redundancy rules are also called 'meta-rules', 'meta-templates', or 'meta-identity-rules' (Janda and Joseph 1986, 1992, 1999) because they identify formal properties that are shared either across rules, or across both rules and representations. In Arapesh these rules do more than express generalizations; they are also necessary in order for class-assignment to proceed as it does, since what is required for a noun to be treated as regular from the point of view of class assignment is that it adhere to such a rule, regardless of whether the rule applies productively to derive the plural, or whether it serves as a redundancy rule merely parsing an irregular and hence necessarily listed form.

The greatest irregularity in plural formation is found in nouns belonging to the semantic classes IV and VII and their cognates. But these are precisely the cases in which it could be argued that class assignment makes reference to something other than morphological rules, since a common semantic feature is available. The most extreme irregularity is found in the class of nouns referring to male humans shown in (27) and (28).

(27) Abu'	<u>Singular</u>	<u>Plural</u>	
	<i>aleman</i>	<i>alemam</i>	'man'
	<i>anen</i>	<i>amum</i>	'husband'
	<i>mohun</i>	<i>mohulim</i>	'male in-law'
	<i>sahalina</i>	<i>sahalihim</i>	'mother's brother'
	<i>ahanina</i>	<i>ahalihim</i>	'father's brother'
	<i>baah</i>	<i>belhehim</i>	'grandfather', 'grandson'
	<i>pater</i>	<i>paterimi</i>	'priest' (TP)
	<i>tisa</i>	<i>tisaimi</i>	'teacher' (TP)

(28) Arapesh	<u>Singular</u>	<u>Plural</u>	
	<i>jaken</i>	<i>jakenim</i>	'father'
	<i>raminen</i>	<i>raheim</i>	'husband'
	<i>araman</i>	<i>aramum</i>	'man'
	<i>awanin</i>	<i>arahim</i>	'younger brother to man'
	<i>barahan</i>	<i>barahowhim</i>	'grandson'
	<i>nigamin</i>	<i>nigamin</i>	'son', 'son in-law'
	<i>arakohonin</i>	<i>arakohonim</i>	'man neutral in warfare'
	<i>megan</i>	<i>meganomwi</i>	'brother in-law to man'

Although the singular forms nearly always end in *n* and the plural forms always contain *m*, there is no way of collapsing the plurals in this class into a unitary generalization with sufficient specificity that it can apply productively, even with the aid of readjustment rules. Nevertheless, the fact that these nouns are classified uniformly in each dialect can be connected with a shared semantic property, male human reference, rather than with adherence to the shared formal generalization $n \leftrightarrow m_{[+PL]}$.

No such argument can be made for Arapesh class IV, however, because it includes not only nouns referring to human females but also many other nouns without human reference.¹¹ As illustrated in (29) below (and discussed further in Chapter 4), Arapesh class IV involves a great variety of irreducible canonical plural endings.

(29) <u>Singular</u>	<u>Plural</u>	
<i>pupijak^u</i>	<i>pupijameb</i>	‘species of insect’
<i>boruboruk^u</i>	<i>boruborumeb</i>	‘crow’, ‘species of tree’
<i>yahak^u</i>	<i>yaharib</i>	‘fruit tree’
<i>muguruk^u</i>	<i>mugurib</i>	‘ear lobe’
<i>unuk^u</i>	<i>unib</i>	‘star’, ‘ogress’, ‘anus’
<i>seraifuk^u</i>	<i>seraifib</i>	‘turtle’
<i>anik^u</i>	<i>aniguhijer</i>	‘species of rattan vine’
<i>mabitek^u</i>	<i>mabiteguhijer</i>	‘duck’
<i>uruwhik^u</i>	<i>uruwhijer</i>	‘side post of a house’
<i>batauifik^u</i>	<i>batauifijer</i>	‘female child’
<i>wharok^u</i>	<i>wharou</i>	‘corn’
<i>abera^u</i>	<i>aberau</i>	‘coconut shell cup’
<i>babwek^u</i>	<i>babwekomi</i>	‘grandmother’
<i>amakek^u</i>	<i>amakekomi</i>	‘mother’s sister’
<i>meganukik^u</i>	<i>meganubiheu</i>	‘husband’s clan brother’s wife’
<i>negauik^u</i>	<i>negaliheu</i>	‘infant daughter’, ‘daughter-in-law’

This class also presents a high degree of stem-internal irregularity, due to intervocalic change and other idiosyncratic alternations (30).

¹¹Irregularity in the Abu’ female human class is much more constrained: nouns with feminine reference take *-wa*, *-iwa*, or *-liwa* in the plural, though which one a given noun takes is unpredictable (Nekitel 1986).

(30)	<u>Singular</u>	<u>Plural</u>	
	<i>ʃekaijok^u</i>	<i>ʃekaimɛb</i>	‘insect with red abdomen’, ‘girl at menarche’
	<i>natemok^u</i>	<i>natepoguhijɛr</i>	‘small snake’
	<i>karahok^u</i>	<i>kararuhijɛr</i>	‘canoe’
	<i>ñenok^u</i>	<i>ñenibijɛr</i>	‘species of snake’
	<i>mohokwik^u</i>	<i>mowhijeliu</i>	‘sister to a man’
	<i>irohokwik^u</i>	<i>ileuliheu</i>	‘wife’

It is not possible to attribute the morphosyntactic unity of these nouns to a common semantic feature. Rather, what unites these nouns in the same morphosyntactic category is the schematic redundancy rule $k^u \leftrightarrow +\text{LABIAL}_{[+PL]}$ that will be discussed in the next chapter.

In the discussion of Arapesh defaults in Chapter 5, I devote some attention to the problem of the default plural suffix *-ehas*. We will see that the lexical presence of this ending serves a blocking function with respect to the regular pluralization rules, while elsewhere this same ending is assigned productively as the default plural to nouns for which no regular plural rule is eligible to apply. In this situation, too, a rule of the form $X \leftrightarrow Xehas_{[+PL]}$ can be called upon to express the notion that *-ehas* instantiates the same plural form whether it is underlyingly associated with a noun lexeme or results from the productive application of a realization rule.

Redundancy rules like the ones appealed to here to parse regularities in the lexicon have a curious status in the history of morphological theory. Despite early arguments pointing to their importance (e.g. Stanley 1967, Jackendoff 1975), redundancy rules are rarely held up as an attractive (or even valid) means of expressing regular linguistic patterns (Bochner 1993 and Booij 2010 are notable exceptions). Still, rules with an explicitly parsing function are often resorted to in individual analyses, many times as footnotes. For example, in Perlmutter’s analysis of Yiddish summarized above, the subset of plurals ending in *-im* are listed in their entirety. The obvious generalization that all these nouns end in the historically Hebrew masculine plural suffix would go unrecognized if Perlmutter did not also posit a redundancy rule: ‘The fact that all [the listed] forms end in *-im* can be stated by a lexical rule’ (1988:91 n.11). Anderson assumes that ‘rules, operating within the lexicon, can describe the systematic relations among stems in the case of irregular stem-alternation patterns, without requiring us to claim that these cases are not distinct from fully suppletive listing’ (Anderson 1982:608 n.15). Redundancy statements have been referred to by various names: ‘via rules’ (Hooper 1976), ‘morpholexical rules’ (Lieber 1990), ‘lexical correspondence rules’ (López 1979), which express ‘non-productive but recurrent alternation[s]’ and ‘bi-directional, non-governed,

non-derivational relationship[s] between underlying segments' (López 1979, cited in Platt 1981:59), 'meta-rules' (Janda 1982, Janda and Joseph 1992), 'pattern matching evaluation' (Bochner 1993), and morphological 'constructions' or 'schemas' (Booij 2010). If related rules or forms are uncollapsible because of their idiosyncrasies, this need not undermine a unified statement of their relatedness; instead, the idiosyncrasies simply constitute additional information that does not contradict the larger generalization. Thus, even when no rule can state the whole truth, we at least can state the partial truths by means of redundancy rules.

Still, there have been few strong empirical arguments in favor of positing redundancy rules on structural grounds; rather, they are typically called upon simply as a convenient descriptive device to prevent generalizations from going unexpressed. But in Arapesh there seems to be no way around them: redundancy rules unifying the plurals of both lexically listed and regularly inflected forms are the basic grammatical structures upon which the entire system of noun classification depends. In the next chapter, we look in more detail at the Arapesh plural categories, focusing on their internal structure.

Organization Within and Across Noun Classes

4.0 Introduction

If we treat class as a category of information associated with noun lexemes, it is sensible to ask how each class is structured internally and how the category as a whole is distributed over nouns in the lexicon. In Section 4.1, I will show that the Arapesh form-based classes have an internal structure that resembles semantically-based categories in being structured radially around a prototype, suggesting that categories of both types are governed by similar cognitive principles.

In Section 4.2, I consider the distribution of classes across nouns. In Arapesh, noun class assignment rules can be distinguished into qualitatively distinct types: semantic, morphological, and phonological. While each type has the power to determine class in certain cases, some rule types take priority over others. The rules themselves are language specific, but the hierarchical arrangement of rule types appears to be a systematic property of language that cannot be derived from general principles of rule interaction.

4.1 The Internal Structure of Noun Class Categories

The literature on noun classification as a kind of categorization has tended to focus on the semantic as opposed to formal bases of nominal categorization systems, both for noun classes and classifier systems (see, e.g., Berlin 1968, Adams and Conklin 1973, Becker 1975, Denny 1976, T'sou 1976,

Allan 1977, Dixon 1982, Craig 1986, Silverstein 1986, Lakoff 1987). One of the ways in which this body of research has converged with categorization research outside of linguistics, especially with the literature on basic level categories (e.g. Rosch and Mervis 1975, Rosch 1978) and cross-cultural research on color categorization (e.g. Berlin and Kay 1969, Kay and McDaniel 1978), is in the recognition that not all members are equally representative of the categories to which they belong. Instead, categories of natural and cultural objects are commonly organized around a prototype, a central or most typical member to which other members are related to varying degrees within a *radial* category (Lakoff 1987:91). In radial categories,

there is a *central* subcategory...; in addition there are *non-central extensions* which are not specialized instances of the central subcategory, but rather are variants of it... These variants are not generated from the central model by general rules; instead they are extended by convention and must be learned one by one. But the extensions are by no means random. The central model determines the possibilities for extensions....

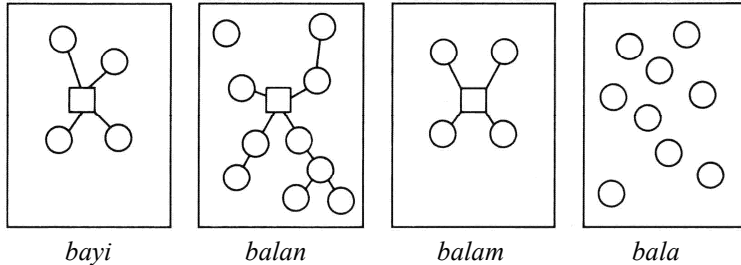
Radial categories can be illustrated with the classifier system of Dyrirbal (Dixon 1982). In Dyrirbal each noun in an utterance must be preceded by one of the four classifiers *bayi*, *balan*, *balam*, *bala*, and though there are semantic resemblances among some of the members of the classes referenced by each of these classifiers, they are neither always obvious nor thoroughgoing.¹ Dyrirbal is not unusual in this regard. As Dixon 1982:178 notes, ‘in most languages which show systems of noun classes there is some degree of semantic correspondence; it is seldom so slight that it can be ignored and seldom sufficient for categorical statements of semantic content.’ So, for example, the Dyrirbal *bayi* class includes men, kangaroos, most fish, snakes and insects, and some spears. The *balan* class includes women, some animals, some spears, some trees, the sun and stars, and things having to do with fire and water. A similar approach to the semantic organization of noun classes in Bantu languages is developed in Spitulnik 1989 for ChiBemba and Contini-Morava 2002 for Swahili. See Aikhenvald 2000:308-311 for further discussion of prototype approaches to noun class semantics.

Lakoff 1987 proposes that the Dyrirbal categories are semantically coherent, though not on the classical view of category membership based on necessary and sufficient conditions. Rather, the Dyrirbal classifiers define coherent radial categories, structured around a central member, with other

¹The Dyrirbal system actually involves four classifier paradigms which encode not only a noun’s class but also its location/visibility and case (Dixon 1982:161). Following common practice, the four absolutive forms built on the stem *bala-* ‘there (and visible)’ are used here to represent these paradigms.

members related by conventional but motivated extensions. The diagram summarizing Lakoff's analysis is reproduced in (1); the basic subcategories to which each classifier corresponds are listed beneath it in (2). The box at the center of the *bayi*, *balan*, and *balam* models represents the central subcategory. The *bala* category lacks such a privileged subcategory, making it the system's default.

(1)



- (2)
- | | | | |
|--------------------------------|--|-----------------------|--|
| <p>human males
animals</p> | <p>human females
water
fire
fighting</p> | <p>non-flesh food</p> | <p>everything else
(i.e., default)</p> |
|--------------------------------|--|-----------------------|--|

Of these, the most interestingly structured radial category is the one defined by the classifier *balan*. Lakoff suggests that the central subcategory of the *balan* category is human females, with a very general 'domain of experience' principle motivating its extension by a process of chaining to other subcategories. The important areas of experience that are relevant to inclusion in the *balan* category are myths and beliefs, fire, and danger. For example, the sun is believed to be the wife of the moon, and so the sun is categorized with human females. Since fire is in the same domain of experience as the sun, fire is categorized with the sun. Similarly categorized are things having to do with fire, including water, which extinguishes fire, and dangerous or stinging things, which burn like fire.

The Dyrbal classifier categories are reanalyzed by Plaster and Polinsky 2010 as partly form-based, with a set of noun-initial phonological predictors (*bi-*, *gugu-*, *ma-*, *yi-*) for the animate nouns that are not accounted for by the main semantic rules referencing sex, edibility, fresh water, fire and stinging. The analysis is helpful for explaining how the classes were traditionally learned; it also makes natural the transformation of the system by the next generation of semi-speakers described by Schmidt 1985 (as we will see, the Weri noun class systems exhibits a similar pattern of loss).

However, it is hard to say whether these formal predictors represent more than implicit patterns in the lexicon. Did they previously serve as productive principles governing linguistic categorization? The one example that could bear on this question in the authors' data is the English borrowing *bigi* 'pig', which is assigned to the *balan* class alongside stinging things, rather than to the *bayi* class, as expected for non-human animates. Assignment to *balan* may indeed follow directly from the noun's initial *bi* sounds. But assignment to *balan* is also suggested by semantic principles, given the assimilation of *bigi* to the same taxonomic group as the echidna in the mother-in-law register, apparently due to the resemblance between the pig's bristles and the echidna's spines (Dixon 1972:311).

In any event, the Arapesh form-based class categories bear a strong structural resemblance to the Dyirbal classifier categories diagrammed above. As we have said, in order to predict the class of an Arapesh noun we need to refer simultaneously to both its singular and plural form; in other words, plural realization plays a critical role in defining the Arapesh classes. Recall as well that most of the classes exhibit some degree of variation in the form of the plural that counts for purposes of inclusion in the class, with one plural form either frequently or formally more typical than the others. For purposes of comparison with the Dyirbal semantic categories, I will focus first on two representative Arapesh classes, class IX and class IV, whose properties are summarized in (3) and (4). The parenthesized numbers appearing next to each of the noun plurals represents the number of times each type is attested in Fortune's 1942 grammar and texts.

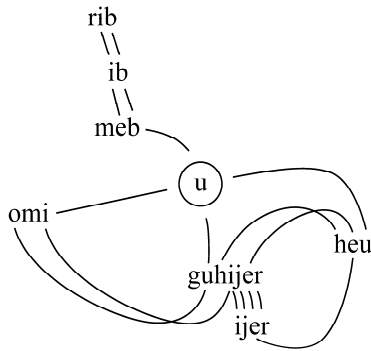
(3)	<u>Class IX</u>	<u>Noun Terminal</u>	<u>Verb Subject Prefix</u>	<u>Possessed Suffix</u>
	Singular	<i>p</i>	<i>p-</i>	<i>-ip</i>
	Plural	<i>s</i> (45)	<i>s-</i>	<i>-is</i>
		<i>gwis</i> (6)		

(4)	<u>Class IV</u>	<u>Noun Terminal</u>	<u>Verb Subject Prefix</u>	<u>Possessed Suffix</u>
	Singular	<i>k^u</i>	<i>kw-</i>	<i>-ik^u</i>
	Plural	<i>meb</i> (14)	<i>w-</i>	<i>-iu</i>
		<i>u</i> (11)		
		<i>ijer</i> (7)		
		<i>guhijer</i> (6)		
		<i>rib</i> (4)		
		<i>heu</i> (4)		
		<i>omi</i> (3)		
		<i>ib</i> (2)		

In the case of class IX, ‘a small subclass drop the *p* of the singular termination and substitute *gwis* to form the plural... [while] the regular form is the change of the singular termination *p* to the plural termination *s*’ (e.g., *bakenop* ~ *bakenogwis* ‘individually owned piece of bush ground’, vs. *agøp* ~ *agøs* ‘species of fruit bearing tree’; see Fortune 1942:31). Analogous asymmetries occur in around half the Arapesh classes.

Class IV exhibits more plural diversity, with no plural standing out as significantly more common than the others. Nevertheless, when the set of plural forms is schematized as in (5), the central subcategory and principle motivating the extensions become clear: the class IV plurals are organized around the feature +LABIAL. This feature is realized most directly in the plural form *u*, which is accordingly the element used in agreement with class IV plurals (alternating with non-syllabic *w* when the segment occurs in syllable-initial position). While it is not possible to predict which forms will constitute class IV plurals, those that are attested are nevertheless motivated, either directly by virtue of a +LABIAL specification in their representations, or else indirectly, by virtue of other formal features they share with category members that are themselves directly motivated. So the plural *-ijer* is indirectly motivated by its formal similarity to *-guhier*, which is itself directly motivated by the labial vowel *u*.

(5)

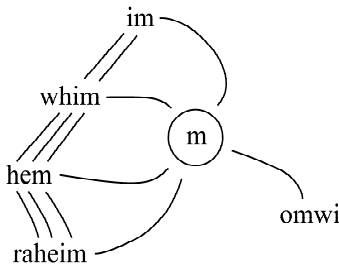


A similar analysis can be given for class VII, the Arapesh male human class. The examples listed in (6) represent the major plural endings found in Fortune’s grammar for nouns of this class.

(6) Singular	Plural	Final Alternation	
<i>araman</i>	<i>aramum</i>	$n \sim m + \text{vowel change}$	'man'
<i>jaken</i>	<i>jakenim</i>	$n \sim im$	'father'
<i>raminen</i>	<i>raheim</i>	$n \sim m + IV\Delta (?)$	'husband'
<i>arakohonin</i>	<i>arakohonim</i>	$n \sim m$	'neutral in war'
<i>megan</i>	<i>meganomwi</i>	$n \sim omwi$	'brother-in-law to a male'
<i>sigaben</i>	<i>sigabehem</i>	$n \sim hem$	'male ghost'
<i>barahan</i>	<i>barahowhim</i>	$n \sim whim$	'grandson'
<i>niganin</i>	<i>nigamin,</i> <i>nigamim</i>	$IV\Delta (n \sim m),$ $n \sim m + IV\Delta (n \sim m)$	'son', 'son-in-law'
<i>mufupimin</i>	<i>mufuhem</i>	$-pimin \sim hem$	'man from Mushu'
<i>walipepimin</i>	<i>walipem</i>	$-pimin \sim m$	'man from Walis'

It is impossible to identify a single suffix that attaches to these nouns in the plural. The intervocalic change type in fact subsumes many more specific patterns of plural formation, since the nature of the internal change varies unpredictably from noun to noun (contrast *aramatokwin* \sim *aramagowem* 'effeminate man, male wastrel' with *awanin* \sim *arahim* 'younger brother to a man'). There are nearly as many plurals as there are nouns in this class. Nevertheless, the formal variation is constrained: every class VII plural form contains the bilabial nasal *m*. This segment is thus the central subcategory of the class VII plural category, the one that motivates the inclusion of more peripheral members like *whim* and even effectively suppletive members like *raheim* (7).

(7)



The Arapesh class IV and VII categories diagrammed in (5) and (7) resemble the sprawling Dyirbal *balan* class diagrammed in (1) in being structured around a central member by motivated but unpredictable extensions. We thus observe a unified principle of categorization holding in both systems of noun classification, regardless of the fact that the defining features

of the category are primarily conceptual in Dyirbal, but primarily formal in Arapesh.

The claim that formally based linguistic categories may have a radial structure is well justified. In their now-classic study of prototype effects in morphology, Bybee and Moder 1983 show that the most productive class of irregular verbs in English (the class whose members include e.g. *spin* ~ *spun*, *cling* ~ *clung*, *stick* ~ *stuck*, *dig* ~ *dug*, *slink* ~ *slunk*) exhibits the structure of a natural category or schema: a partially specified prototypical form to which less prototypical category members are associated by degrees of family resemblance. The English irregular past tense schema has the basic structure s(C)(C)ɪŋ ~ s(C)(C)ʌŋ, with the final velar nasal serving as the single strongest predictor of whether a nonce verb will be inflected according to the irregular pattern.

Köpcke 1988 extends this argument to the entire system of plural marking in German. Although the German plural category uses a set of formally diverse markers, Köpcke has argued that it is schema-driven, with formal features of the singular influencing which plural form a noun will select. The continuum model Köpcke arrives at specifies both an idealized singular and an idealized plural form, which can be construed as the prototypes for their respective number categories. A prototypical singular noun in German is monosyllabic, ends in a stop consonant, and takes either *der* or *das* for its determiner. A prototypical German plural is polysyllabic, ends with the segment(s) -(ə)n, and takes *die* for its NOM/ACC determiner.

Obviously, these prototypes do not determine plural form in any given case; nor do they recognize every generalization there is to make about the system, e.g., that nouns ending in certain derivational suffixes and feminine nouns ending in schwa 'reach near categorical assignment of a plural suffix based on the singular form of the noun' (Köpcke 1988:330). But the notion of schema addresses the same descriptive challenge for which Lakoff invokes 'motivation' in accounting for the structure of semantic categories: the need for a device that transcends the tension between the two poles of arbitrary listing vs. full conformity to a rule. That is not to say that there are no lists of irregular verbs in the English lexicon or rules of plural formation in German morphology. What the notion of schema implies, rather, is that the application of rules to a given input form or the selection of a form from a list is guided at least in part by an idealized knowledge structure. Given the fact that his participants 'showed marked deviations from real lexical patterns', Köpcke concludes (1988:319-320) that

a model based on major patterns [i.e. the 'members' of the category of German plurals] in combination with the principle of cue strength [i.e., features such as perceptual salience and frequency which jointly influence which plural markers are favored] suggests that speakers' mental repre-

sentation of morphological knowledge simplifies immanent [formal] patterns in the lexicon in accordance with general cognitive principles.

The formal schemas proposed by Köpcke for German nouns and by Bybee and Moder for English irregular verbs both associate phonological shape with morphological categories, much as Lakoff's analysis of the Dyrbal *balan* category associates a conceptual category with a particular morphological category (a noun class). But there is a linguistic reality to formally-defined schemas even when they fail to instantiate a unified morphological category. Janda 1982 argues that the high frequency with which morphologically distinct non-root elements exhibit identities, or near-identities that simultaneously prevent their collapse into unitary generalizations, suggests that a theoretically anomalous preference for morphological homophony is at work in natural language. In a standard rule-based system it is impossible to express these kinds of generalizations, which are 'found' in the data yet 'lost' to explanation. Such patterns can be accounted for by positing polysemous or asemous super-structures (meta-templates, meta-identity rules, or morphological rule constellations) that operate by parsing the various occurrences of the elements that appear in lexical entries and morphological rules, where they are connected with specific meanings or functions and are given concrete expression. An example is the aorist category of the verb in Sanskrit. There are no less than eight distinct patterns of prefixal reduplication used to mark this category, none of which can be derived by general phonological principles: some (but not all) contain a pre-specified vowel; some (but not all) induce changes in the stem; the reduplicated prefixes vary in their phonological weight. Nevertheless, all these diverse prefixal formations—in addition to still others that are associated with yet other morphological values—can be unified by a meta-template of the form [...V...][ROOT], which 'expresses the constant element that is shared across the entire set of reduplicative prefix-templates' (Janda and Joseph 1992:167). The Sanskrit meta-template obviates neither the need to associate actual reduplicative patterns or templates with the categories they mark nor the need to list them in the lexical entries of most words. But it does identify the central tendency of the formal element 'reduplicative prefix' in the language. In effect, the meta-template functions analogously to the central subcategory of a radial category.

The schematic noun class structures diagrammed in (5) and (7) above similarly lack a unified signaling function. In order to see how this is the case, we need to look more closely at the phenomenon of intervocalic change, whereby alternations that are associated with a particular class when they appear at the ends of nouns occur noun-internally, on a lexical basis, with no bearing on the noun's morphosyntactic categorization. Inter-

vocalic change (here symbolized IV Δ) occurs in every variety of Arapesh. It was originally described by Fortune 1942 and is as ubiquitous a pattern in Arapesh nominal morphology as the system of noun classes itself. Some examples from languages across the family are given in (8). The alternating segments are indicated in the IV Δ column, where the Roman numeral in parentheses indicates the Arapesh class referenced by the word-internal alternation.

(8)

	<u>Singular</u>	<u>Plural</u>	<u>IVΔ</u>	
Balam	<i>balhanin</i>	<i>balhamim</i>	<i>n ~ m</i> (VII)	‘grandson’
	<i>fulukum</i>	<i>fuluwep</i>	<i>k ~ w</i> (IV)	‘fruit’
Hwamsək Abu’	<i>kwafita</i>	<i>kwasitok</i>	<i>f ~ s</i> (IX)	‘spoon’
	<i>hambomuna</i>	<i>hambufinab</i>	<i>m ~ f</i> (V)	‘owl’
Weri	<i>bipitenger</i>	<i>bistonguh</i>	<i>p ~ s</i> (IX)	‘bone’
	<i>amborip</i>	<i>ambohwis</i>	<i>f ~ h</i> (X)	‘shadow’
Mufian	<i>awinil</i>	<i>awisingof</i>	<i>n ~ s</i> (VIII)	‘centipede’
	<i>bangata</i>	<i>bahatangw</i>	<i>ng ~ h</i> (III)	‘sago leaf rib’
Bukiyip	<i>alatihw</i>	<i>alagwilih</i>	<i>t ~ gw</i> (XI)	‘feast’, ‘party’
	<i>niganini</i>	<i>nigamim</i>	<i>n ~ m</i> (VII)	‘son’

As with the choice among competing plurals, the occurrence of these noun-internal alternations is not predictable. Nevertheless, when they do occur, they typically repeat the patterns also found at the ends of nouns. The following survey of a large semantic domain in one dialect provides a sense of how thoroughgoing this pattern of lexical alternation is.²

Eighty-eight nouns naming kinds of birds were recorded for Cemaun Arapesh. The nouns range over ten different classes. Of these eighty-eight nouns, seventy-seven have plurals that are regular for purposes of form-based class assignment, even if they are not strictly predictable. Of the eleven remaining bird nouns, nine are only irregular in that their plurals

²The ubiquity of intervocalic change across classes and semantic domains in all Arapesh varieties is consistent with the general exuberance of plural marking we find in these languages. For example, in an unpublished Bukiyip text (Conrad n.d.:HN12), the borrowed Tok Pisin noun *tebol* ‘table’ is pluralized as *teleboguhwas* with no fewer than three partially overlapping markers: the class II metathetic plural SG *bVI* ~ PL *IVb*, the class X plural *guh* that applies to liquid-final nouns, and the default class VIII plural *-ehas*. Evidently, speakers feel the pressure described by Thomason 1988 to keep adding affixes to morphologically complex forms, resulting in semiproductive multiple exponence.

contain intervocalic changes. These correspond to four distinct classes as shown in (9a). The two further bird nouns in (9b) have plurals that are irregular in other ways.

(9)	<u>Singular</u>	<u>Plural</u>	<u>IVΔ</u>	
a.	<i>mañin</i>	<i>macib</i>	<i>ñ ~ c</i>	‘pigeon’, ‘dove’
	<i>wiñin</i>	<i>wicib</i>		‘Victoria crowned pigeon’
	<i>cepleñip</i>	<i>ceplecis</i>		‘papuan flowerpecker’
	<i>gätäpen</i>	<i>gätäsenəb</i>	<i>p ~ s</i>	‘species of bird’
	<i>woripak^w</i>	<i>worisau</i>		‘coastal wildfowl’
	<i>kwipun</i>	<i>kwisib</i>		‘species of bird’
	<i>yurupin</i>	<i>yurusirəb</i>		‘frog-faced bird’ ³
	<i>səbiteñ</i>	<i>səbigwec</i>	<i>t ~ g^w</i>	‘honey-eater’
	<i>yəharik^w</i>	<i>yəhaguhiu</i>	<i>r ~ guh</i>	‘species of bird’
b.	<i>biok^w</i>	<i>biorumeb</i>	—	‘species of bird’
	<i>bogerik^w</i>	<i>bogeribihər</i>	—	‘nightjar’

The phenomenon of intervocalic change probably traces back to the Arapesh protolanguage since there are instances of cognate intervocalic changes in items of core vocabulary (the word ‘bone’ is *bələpīgər ~ bəsigu* in Rohwim, *biripiger ~ birisigu* in Cemaun, and *biḫtakel ~ bistakuh* in Hwamisək Abu³). At the same time, these patterns in the lexicon can in some cases be activated synchronically. For one thing, some Tok Pisin borrowings have something approaching intervocalic change at the junction of the stem and the default plural suffix, e.g., ‘overapplications’ like Cemaun *kap ~ kasəhəs* ‘cup’ and Hwamsək Abu’ *kiap ~ kiasimi* ‘kiap, patrol officer’. For another, speakers occasionally ‘err’ in producing intervocalic change plurals that others (or they themselves) will later reject, especially for uncommon nouns. Given the Arapesh-wide tendency for nearly every noun-final sound to have a distinct corresponding plural, for sounds to give way to their plural alternants in a plural context—especially near the ends of words—is certainly well-motivated by the patterns in the lexicon.

It is sometimes possible to interpret intervocalic change nouns as derived through a syntactic process of possessive formation as shown in (10) (data from Fortune 1942), because both the semantics and the form are transparent. The word for ‘*agəp* leaves’ in (10c), for example, is essentially

³This noun has an alternate plural shown in (14).

the phrase ‘*agøp*-POSS-leaves’ with the possessed noun ‘leaves’ unexpressed.

(10)	<u>Singular</u>	<u>Plural</u>	<u>IVΔ</u>	
a.	<i>aragokwil</i>	<i>aragamebiguh</i>	kw ~ meb (IV)	‘sow’
	<i>aragok^u</i>	<i>aragomeb</i>		‘female’
	<i>bul</i>	<i>buruguh</i>		‘pig’
	<i>aragokw-i-l̄ (bul)</i>			
	female-POSS-AGR (pig)			
b.	<i>aramatokwiñ</i>	<i>aramogowef</i>	kw ~ w (IV)	‘female infant’
	<i>aramatok^u</i>	<i>aramogou</i>		‘woman’
	<i>batauiñ</i>	<i>batauiñ</i>		‘child’
	<i>aramatokw-i-ñ̄ (batauiñ)</i>			
	woman-POSS-AGR (child)			
c.	<i>agøpip</i>	<i>agøsis</i>	p ~ s (IX)	‘ <i>agøp</i> leaf’
	<i>agøp</i>	<i>agøp</i>		‘variety of tree’
	<i>fup</i>	<i>fus</i>		‘leaf’
	<i>agøp-i-s̄ (fus)</i>			
	trees-POSS-AGR (leaf)			

But whatever role such constructions may have played in the historical development of intervocalic change nouns, very few are analyzable in this way synchronically. Scanning down the list of bird names in (9), for example, were we to extract everything to the right of the intervocalic changes, we would not be left with independent nouns or even identifiable stems of any part of speech; we would simply be left with fragments of the words.

Moreover, for some cases of intervocalic change in which a possessive structure is defensible on formal grounds, the corresponding semantic relationship does not obtain, undermining a constructional analysis as in (10). For example, the Cemaun Arapesh noun SG *əgulañig* ~ PL *əgulacigəs* ‘type of tree’ appears to have the possessive structure SG *əgulañ-i-g* ~ PL *əgulac-i-gəs* given the existence of the independent noun SG *əgulañ* ~ PL *əgulac* ‘type of vine’. Both are jungle materials used (among other things) in house-building, but speakers do not recognize a morphological relation between the two nouns; *əgulac* vines do not grow on *əgulacigəs* trees nor

resemble them in any way that speakers are conscious of.⁴ Nor is it clear what implied noun the apparent class-marking final SG *g* ~ PL *gəs* would correspond to, since nouns denoting trees can fall into any of eleven noun classes.

Finally, there are cases for which a syntactic analysis of intervocalic change is formally untenable, because the alternation is embedded too far to the left of the word. The nouns in (11) are examples. In each case, a non-alternating segment intervenes between the intervocalic change and the noun-final alternation: *b*, *n*, and *w*, respectively.

(11)	<u>Singular</u>	<u>Plural</u>	<u>IVA</u>	
	<i>yawiñibok^w</i>	<i>yawicibou</i>	<i>ñ ~ c</i>	‘type of tree’
	<i>əbetənəh</i>	<i>əbegwoneh</i>	<i>t ~ g^w</i>	‘thumb’, ‘big toe’
	<i>areñiwog</i>	<i>arecuwos</i>	<i>ñ ~ c</i>	‘wild sugar cane leaf’

Something along the lines of intervocalic change occurs when an individuating suffix is added to a numberless noun, as if deriving a new singular form in the process (12a). The more common pattern adds the individuating suffix directly to the noun’s base form (12b). In these examples, all of which refer to mushrooms, the final *p* has a classifier-like function with its apparent source in the noun *cup* ‘leaf’.

(12)	<u>Base Noun</u>	<u>Individuated Form</u>	<u>‘mushroom’</u>
a.	<i>araguh</i>	<i>ararip</i> (* <i>arar</i>)	generic term
	<i>kudalic</i>	<i>kudaliñip</i> (* <i>kudaliñ</i>)	jellylike variety
b.	<i>ubibiguh</i>	<i>ubibiguhip</i>	red variety
	<i>apoguh</i>	<i>apoguhip</i>	white variety
	<i>səsaweh</i>	<i>səsawehip</i>	grows on sago palm rib
	<i>cukwehəbir</i>	<i>cukwehəbirip</i>	grows on sago pith

Intervocalic change shows up in other surprising places. It occurs in at least one adjective stem: SG *kwahapi-neri* ~ PL *kwahasi-mi* ‘large’. It occurs in two Weri kinship nouns that are inflected lexically for their possessor, SG *bahrumi-ə* ~ PL *bahrupi-ə* ‘male.in.law-my’ and SG *bahrənom-iə* ~ PL *bahrəp-iə* ‘grandfather-my’ (note that the class indicated by the intervocalic

⁴Similarly, Cemaun speakers rejected the suggestion that *urupat* ~ *urusag^w* ‘house’ might bear some non-accidental relation to *urup* ~ *urus* ‘grass skirt’. Hwamsək Abu’ speakers laughed when I observed that *dubarin* ~ *dubarub* ‘hornbill’ is embedded in *dubarinikil* ~ *dubarubukuh^w* ‘four-corner bean’. It is not that the hornbill eats that kind of bean or resembles it in some way. Instead, they said, ‘God just gave those things names that sound the same.’

change, $m \sim p$, is not the one expected for male humans in Weri, $n \sim h$). The Cemaun Arapesh noun *apotoh^w* ~ *asogurərih* ‘boil, abscess’ has two unanalyzable intervocalic changes ($p \sim s$, $t \sim gw$), as does *amapokwig* ~ *amasowigəs* ‘species of tree used for making posts’ ($p \sim s$, $kw \sim w$).⁵ In the Cemaun noun *mədīdānih* ~ *mədīdābih* ‘variety of short yam’ the intervocalic change ($n \sim b$) is the only alternation differentiating the singular form from the plural. And the base of the Rohwim derived form *Walipepimin* ‘man from Wallis’ is back-formed by ‘undoing’ an intervocalic change in the final consonant of the otherwise non-alternating proper name *Wallis* (Fortune 1942:25).

In (13) we see that the constrained irregularity that is so common in Arapesh pluralization can extend even into the intervocalic changes. Thus, not only is it unpredictable whether or not intervocalic change will occur in a given noun, but even the *form* of the intervocalic change is unpredictable while yet falling within the range of formal variation exhibited noun-finally. The alternations here are independently attested members of the radially structured class IV plural.

(13) <u>Singular</u>	<u>Plural</u>	<u>IVΔ</u>	
<i>jakerok^u</i>	<i>jaberemeb</i>	$k \sim b$	‘tail of a tree kangaroo’
<i>meganukik^u</i>	<i>meganubiheu</i>	$k \sim b$	‘sister-in-law to a woman’
<i>burukwiuh</i>	<i>buruweruh</i>	$kw \sim w$	‘village or locality section’
<i>irukweuh</i>	<i>iruweruh</i>	$kw \sim w$	‘coconut husk’
<i>mohokwik^u</i>	<i>mowhijeliu</i>	$kw \sim ijel$	‘sister to a man’

There are three different plural subpatterns that correspond either partially or fully to those found word-finally on class IV nouns: one replacing k^u with b , another replacing k^u with u , realized intervocalically as the glide w , and a third replacing k^u with *ijel* (a variant of word-final *ijer*). In other words, the same class-defining canonical number pairings also appear as lexicalized alternations within words.

In (14) we see the intervocalic changes can even pattern in reverse: the segment normally associated with the singular appears in the middle of the plural form and vice versa. There are not many examples like these, but the fact that we find any underscores how strongly the canonical segmental alternations are associated with number marking in Arapesh, even though they cannot be said to signal number directly.

⁵Speakers deny there any relationship between this noun and *amapok^w* ~ *amasou* ‘frog’.

(14)		<u>Singular</u>	<u>Plural</u>	<u>IVΔ</u>	
	Cemaun	<i>yurupin</i>	<i>yuruminəb</i>	<i>p ~ m</i>	‘frog-faced bird’
	Mufian	<i>lowaf</i>	<i>lu’ongof</i>	<i>w ~ ’</i>	‘clothes’

Despite their obvious formal similarity to the alternations appearing at the ends of nouns, when they occur in the intervocalic environment, these alternations have no influence on a noun’s class. In other words, the patterns themselves, along with all their idiosyncrasies, are detachable from the morphosyntactic categories that they typically signal, in which case they serve at most as a generalized sign of number, again more or less tightly organized around a formal core.

Morphological constellations like these may arise when distinctions are collapsed over time due to phonological erosion, or when the unique signifier of a given formal element is splintered through historical reanalysis. Yet diachronic factors alone are not a fully satisfying explanation for the cross-linguistic prevalence of morphological constellations, since the participating forms are recapitulated in generation after generation of synchronic grammars (Janda 1982). The fact that the same radial category structure appears in both the conceptually based classes of Dyirbal and the primarily form-based classes of Arapesh underscores the generality and flexibility of our capacity to construct such categories, even around form that serves no direct signaling function in a linguistic system.

As an aspect of human experience, formal patterns are available for manipulation and structuring according to the most general psychological mechanisms. Lakoff 1987:331 makes precisely this point in regard to the use of English discourse deictics. In a sentence like *John and Bill came into the room in that order*, the expression *that order* refers not to representations of entities in the world, but rather to the order of the elements *John* and *Bill* in the sentence.⁶ The point is that linguistic elements as such are among the perceptible entities in the world. In the case of nominal classifiers, these perceptible entities include types of human interaction with the objects they classify. Denny 1976 distinguishes these into three subtypes: physical interaction involves the arrangement of objects according to how we manipulate them; functional interaction involves the use to which we put objects; social interaction involves reinforcing distinctions among selected culturally relevant members of society. Extrapolating somewhat from Denny’s typology, it is possible to understand purely formal classification of the sort associated with morphological constellations as a sub-type of physical interaction: recurring formal elements comprise natural targets for

⁶The original observation is due to Ross 1970.

physical categorization in terms of both the motor activities people engage in to produce them and the perceptual activities they use to identify them.

In a sense, then, formal constellations like the Arapesh class IV plural constitute latent grammatical categories, structures that may—but need not—be anchored in consistent morphological meaning. We should even predict the existence of such categories; after all, despite the great emphasis the Saussurean legacy places on an arbitrary relation between signifier and signified, there is no doubt that some sort of structure holds independently on either side of the form-meaning pairing. There is certainly evidence for linguistically significant conceptual categories that are not anchored in linguistic form. No one would question the purely conceptual relation between words *cockatoo* and *robin* in English despite the fact that the formal element /brd/ is not a part of their lexical entries. Priming with semantically related words has repeatedly been shown to facilitate performance in lexical decision tasks, arguing that conceptual structures may be activated independent of their lexical anchors (see Tanenhaus and Lucas 1987:224-226). Accordingly, purely conceptual categories are typically assumed to have a distinct representation in models of lexical processing, for example in the ‘concept nodes’ posited by connectionist models of lexical meaning (see, e.g., Schreuder and Flores D’Arcais 1989:422-426).

To summarize, conceiving the Arapesh classes as categories of information associated with lexemes allows us to understand even the elaborate radial structure of the Arapesh class IV plural category as the product of a general categorizing capacity that creates similar sorts of structures regardless of whether it applies to concepts or to morphological form, including form that cannot be reduced to a specifiable grammatical category.

4.2 Organization Across Noun Class Categories

This section introduces an issue of central importance for the rest of the book, namely, the major types of class-assignment rules that coexist in a morphological system and the principles that decide among them as they compete to divide the set of nouns in a language into morphosyntactic classes. The goal of the present section is simply to show that Arapesh class assignment rules based in nominal semantics take systematic priority over rules based in nominal form. Two major surveys, Corbett 1991 and Aikhenvald 2000, find this pattern of semantic precedence to be a regular property of noun classification systems. We will therefore consider whether it can be attributed to the same general principle of rule interaction that governs other rules in morphology, just as the internal structure of class categories can be attributed to general principles of categorization.

Arapesh class assignment is extraordinary in the degree to which it is sensitive to the phonological shape of nouns. But it is at the same time quite typical in incorporating two clear semantic assignment rules, those classifying nouns denoting male and female humans (class VII and class IV). Examples of Rohwim nouns in these semantic classes are repeated in (15) and (16), respectively.

(15)	<u>Singular</u>	<u>Plural</u>	
	<i>aramatok^u</i>	<i>aramagou</i>	‘woman’
	<i>arapeñik^u</i>	<i>arapefijer</i>	‘female friend’
	<i>babwek^u</i>	<i>babwekomi</i>	‘grandmother’
	<i>fekaijok^u</i>	<i>fekaimeb</i>	‘girl at first menstruation’
	<i>nigauwik^u</i>	<i>nigaliheu</i>	‘daughter’, ‘daughter in law’
	<i>unuk^u</i>	<i>unib</i>	‘mythical woman with teeth in vagina’
(16)	<u>Singular</u>	<u>Plural</u>	
	<i>araman</i>	<i>aramum</i>	‘man’
	<i>afuken</i>	<i>afukenim</i>	‘older brother to man’
	<i>megan</i>	<i>meganomwi</i>	‘brother in law to man’
	<i>niganin</i>	<i>nigamin</i>	‘son’, ‘son in law’
	<i>raminen</i>	<i>raheim</i>	‘husband’
	<i>mufupimin</i>	<i>mufuhem</i>	‘man from Mushu Island’

Even when class is based in semantics, there is a simultaneous correlation with phonological form. Nouns denoting female persons end with the phonological sequence *k^u* in the singular and with a member of the labial schema discussed above in the plural. Nouns denoting male persons end with *n* in the singular and with a member of the *m* schema in the plural. There is thus a systematic overlap between formal and semantic assignment criteria, with the two principles typically converging on a common classification. Similar markers are found in all Arapesh dialects.

(17)		‘man’ (VII)		‘woman’ (IV)	
		<u>Singular</u>	<u>Plural</u>	<u>Singular</u>	<u>Plural</u>
	Bukiyip	<i>əlman</i>	<i>əlmom</i>	<i>əlmatok</i>	<i>əlmagou</i>
	Mufian	<i>aman</i>	<i>amam</i>	<i>nemata^w</i>	<i>nematawa</i>
	Abu’	<i>aleman</i>	<i>alemam</i>	<i>numata[’]</i>	<i>numatawa</i>
	Weri	<i>emen</i>	<i>omoh⁷</i>	<i>opondu[’]</i>	<i>opondowa</i>

⁷The agreement marker for male human plurals is *h^w* in Weri.

The overlap is not complete, however. Class IV also includes many other nouns that lack female-person semantics.

(18)	<u>Singular</u>	<u>Plural</u>	
	<i>binuk^u</i>	<i>binumeb</i>	‘back of neck’
	<i>yahak^u</i>	<i>yaharib</i>	‘a fruit tree’
	<i>seraifuk^u</i>	<i>seraifib</i>	‘turtle’
	<i>urumek^u</i>	<i>urumeguhijer</i>	‘ruff of feathers on bird neck’
	<i>karahok^u</i>	<i>kararuhijer</i>	‘canoe’
	<i>wharok^u</i>	<i>wharou</i>	‘type of indigenous corn’

As is the case with terms referring to women, this semantically diverse set is not restricted as to which of the various forms of the class IV plural they take.⁸ It is thus possible to subsume all class IV nouns under a single rule that makes reference to their *k^u*-final singular form. Aronoff 1994:112 suggests that ‘[f]or female persons... we may posit a derivational suffix *-k^u* [that] will act like other instances of the phonological sequence *k^u*’. This is a reasonable suggestion to the extent that the final *-k^u* of female person nouns alternates in some cases with male human *-n* (e.g., *afukek^u* ‘woman’s elder sister’ ~ *afuken* ‘man’s elder brother’). One disadvantage of the derivational-suffix approach is that it undermines the underlying unity of the terminal element *k^u*, though in a language that already countenances so many redundant representations of the same underlying ‘morpheme’, one more may not be too troubling. A more serious disadvantage is that the purported suffix is totally unproductive: *-k^u* (or its cognates) is never applied to any introduced noun with a female human referent in any Arapesh language. Such nouns are adapted only phonologically with respect to their singular form, and assigned to class IV nonetheless. Furthermore, not all native female person nouns have it, e.g., Cemaun *məgan* ‘sister-in-law’ and hybrids like *nakur* ‘one’s own parent in-law’⁹. So one way or another we need two distinct but overlapping rules, one assigning nouns with female human referents to class IV, and another assigning *k^u*-final nouns to class IV irrespective of their semantics.

As shown by cases like *məgan*, not all *n*-final nouns have male person semantics. Unlike the related division among *k^u*-final nouns, however, the main distinction between masculine and non-masculine *n*-final nouns is

⁸The one exception to this statement is the plural ending *-om(w)i*, which applies exclusively to kinship terms (e.g. Cemaun *babwek^u* ~ *babwekomi* ‘grandmother’, *wawen* ~ *wawenomi* ‘mother’s brother’).

⁹Cemaun plural *nakurehem*. The indirect relation ‘parent-in-law of another’ is differentiated for gender: *nakurikik^w* F, *nakurinen* M.

reflected in class assignment: Arapesh nouns that end in *n* but that do not refer to humans take plurals ending with *b* (e.g., *wai aun* ~ *wai aub* ‘modern bead’, *jehin* ~ *jehinab* ‘tree kangaroo’) and agreement with class VI, distinguished from class VII by the presence of *b* in the plural agreement forms, as opposed to *m*. This lack of isomorphism between class and singular form is uncharacteristic in the language, and in this case it also coincides with the absence of a predictable relation between singular and plural form, since for *n*-final nouns, which of the two possible plural schemas a noun selects depends on semantic properties. Class assignment for *n*-final nouns is thus unusual in a cluster of related respects, and constitutes an interesting exception to the phonological prediction scheme laid out in Chapter 2.

One of the main generalizations to emerge about classification systems is that they always have a semantic core. There are three ways in which this semantic core is manifested: ‘[G]ender always has a basis in semantics. Furthermore, when semantic and formal criteria are both involved in class assignment, they always overlap to some extent’ (Corbett 1991:63). Finally, ‘when different [types of] assignment rules conflict, normally the semantic rules take precedence’ (1991:66). We will consider each of these statements in turn.

No matter how numerous and intricate the formal rules in a given class system may be, they are always accompanied by at least one semantic rule. A related observation was made by Sadock 1983:202-3 in arguing for the existence of overlap among what are in principle independent grammatical generalizations:

[I]t is a universal fact, apparently, that the features that are spread in [agreement...] are never ‘grammatical’ in the strictest sense of the word. Invariably, they have some semantic ramifications.... Indeed, when grammatical classes are totally independent of semantics, as is the case, for example with Latin conjugation and declension classes..., the features defining these classes tend not to be spread by agreement to modifiers’.

Lapointe 1988:83 proposed this as a universal constraint on gender systems: ‘Every language exhibiting gender classes has at least one semantic base.’ In other words, the key defining feature of gender or class, syntactic agreement, depends on the presence of a semantic association somewhere in the system of classes, however narrowly circumscribed that association may be. It is not entirely clear just why this should be the case. It is true even in Arapesh, the great majority of whose classes are semantically arbitrary.

The systematic nature of this distinction among rule types is underscored by the noun class system of Weri. Like other Arapesh varieties, Weri has a large number of classes for realizing number on nouns. Examples of canonical singular/plural pairs are shown in (19), where forms documented in Leavitt nd:a are tagged [SL], and items attested in both Leavitt’s and my

fieldnotes are marked [&SL]. Items documented in my own fieldnotes only are unmarked.

(19) <u>Sg ~ Pl</u>	<u>Singular</u>	<u>Plural</u>	
<i>t(V(h)) ~ ous</i>	<i>kəratəh</i>	<i>kəratous</i>	‘stick’
	<i>wombətə</i>	<i>wombətous</i>	‘slit gong’ [&SL]
	<i>mout</i>	<i>mous</i>	‘ant’ [SL]
<i>r ~ (g)Vh</i>	<i>ember</i>	<i>omboh</i>	‘axe’, ‘water’ [SL]
	<i>susu’ur</i>	<i>susu’uh</i>	‘cockroach’, ‘bedbug’ [SL]
	<i>inimboŋger</i>	<i>inimboŋguh</i>	‘guinea fowl’ [SL]
<i>n ~ s</i>	<i>mahin</i>	<i>mahis</i>	‘meat’ [&SL]
	<i>rehin</i>	<i>rehis</i>	‘stirred sago’ [&SL]
	<i>utumbuen</i>	<i>utumbues</i>	‘wood-eating insect’ [SL]
<i>p ~ s</i>	<i>indap</i>	<i>indas</i>	‘wooden plank’ [&SL]
	<i>suhup</i>	<i>suhis</i>	‘leaf’
	<i>oŋgotop</i>	<i>oŋgotohis</i>	‘knife’ [&SL]
<i>’w ~ wa</i>	<i>aha’w</i>	<i>ahawa</i>	‘coconut’
	<i>itaru’</i>	<i>itarowa</i>	‘praying mantis’ [SL]
	<i>di’u’</i>	<i>di’urahuwa</i>	‘banana’
<i>h^(w) ~ rVh</i>	<i>ruwah</i>	<i>ruwarəh</i>	‘cloth used to carry baby’
	<i>de’wəh</i>	<i>de’worəh</i>	‘stone’, ‘coin’, ‘money’
	<i>ka’imboh</i>	<i>ka’imbəreh</i>	‘snake bean’
<i>mber ~ remp</i>	<i>imbir</i>	<i>irimp</i>	‘stomach’ [SL]
	<i>irember</i>	<i>ireremp</i>	‘ginger’ [SL]
	<i>koti’ember</i>	<i>koti’eremp</i>	‘knock’
<i>m ~ V_[+FR]p</i>	<i>wasum</i>	<i>wasip</i>	‘wild sugar cane’ [&SL]
	<i>apum</i>	<i>apip</i>	‘fruit’
	<i>ma’uhitom</i>	<i>ma’uhitep</i>	‘bee’
<i>ŋg(V) ~ s</i>	<i>muririnəŋgə</i>	<i>muririnəŋgəs</i>	‘large black millipede’
	<i>kumbrəŋga</i>	<i>kumbrəŋgas</i>	‘rat’
	<i>roweŋ</i>	<i>rowəs</i>	‘tree’, ‘wood’
<i>n(V) ~ bVs</i>	<i>waren</i>	<i>warembes</i>	‘species of wildfowl’
	<i>usinə</i>	<i>usinəmbis</i>	‘tree kangaroo’
	<i>suŋgwana</i>	<i>suŋgwanabis</i>	‘walking stick’
<i>mb ~ bVs</i>	<i>səpə’əmb</i>	<i>səpə’əmbəs</i>	‘fingernail’, ‘claw’, ‘hoof’
	<i>oumb</i>	<i>owis</i>	‘coconut’

Pluralization of borrowed nouns in Weri reveals the same characteristic sensitivity to singular-final phonological form we find in the other Arapesh languages: *t*-final *bet* (TP ‘shelf’) is pluralized *betous*, *r*-final *gar* (TP ‘car’) is pluralized *gah^w*, *m*-final *lam* (TP ‘lamp’) has the plural *lap*, and so on, with a high degree of regularity.

But Weri is unique among Arapesh varieties in that all these morphological classes are equivalent for purposes of syntactic agreement.¹⁰ It is still possible to identify constant agreement forms cumulatively indexing class and number that are constant across functions (e.g., cross-referencing both subject on verbs and possessor on nouns) and positions (e.g., used in both prefix and suffix slots), though additional material may sometimes be necessary to signal particular functions (e.g., the adjective endings are *-nəri*, *-hwəri*, etc.). But Weri agreement is distinguished only along semantic lines: there are three classes for humans (male, female, and human other), plus a residue class for nouns referring to anything else (20). Thus all the nouns listed in (19) call for the same set of singular/plural endings on modifying adjectives, possessive pronouns, verbs, etc., regardless of the controlling noun’s form and in contrast to nouns with human referents:¹¹

(20)	<u>Singular</u>	<u>Plural</u>
Human Male	- <i>n</i>	- <i>h^w</i>
Human Female	-’ <i>i</i> ’ <i>w</i>	- <i>w</i>
Human Other	- <i>n</i>	- <i>s</i>
Nonhuman	- <i>n</i>	- <i>m</i>

In Weri, then, we see an Arapesh class system stripped down to a core set of classes referencing precisely the semantic categories upon which formal

¹⁰For the male and female human classes, the agreement markers alliterate with the canonical forms associated with those classes of nouns (or noun stems, in the case of kinship terms), just as in other Arapesh languages. Thus, the plural agreement marker for male humans echoes the *h^w* we find finally on *omoh* ‘men.’ Interestingly, in Cemaun, Rohwim, and Bukiyip, there is only one paradigm in which an agreement marker diverges phonologically from the canonical form of the noun: whereas the plurals of male human nouns take agreement forms marked alliteratively by *m* (to my knowledge the pattern in all Arapesh languages but Weri), the subject marker on verbs with male human referents is *h*. This ‘displacement’ of alliterative agreement is curious in light of the fact that Ross n.d. reconstructs **m* for this paradigmatic position in proto-Torricelli, suggesting that non-alliterative *h^w* is an Arapesh development that reaches its fullest extent in Weri, where *m* marks agreement with nonhuman plurals. It is possible that *h^w* was dragged in to re-distinguish male human nouns for purposes of reference tracking once agreement with all nonhuman plurals collapsed to *m*.

¹¹Leavitt records the nonhuman singular adjective suffix as *-eri*. I assume this represents dialect variation.

assignment is overlaid in the other varieties. It is common to find core semantic class distinctions persisting even when noun classes are in the process of being lost (Demuth, Faraclas, and Marchese 1986)

In order to understand Corbett's second statement that there is always overlap among semantic and formal criteria, imagine a language Arapesh* in which the form-based assignment rules were out of step with the core semantic rules.¹² Nouns referring to female humans would be assigned to class IV, but the residue of *k''*-final nouns with non-female human semantics would be assigned not to class IV, but instead systematically to some other class. Because semantic and formal rules would not lead to the same assignment, knowledge that a noun ends in *k''* would not be a useful clue to its morphosyntactic categorization apart from meaning. Systems like this are not found; as we have seen, there is overlap among rule types even in Arapesh. As Corbett notes, this systematic redundancy is useful to speakers: 'While it may be possible to show that for nouns of a particular type one type of assignment rule takes precedence, and so is generally of greater importance than the others, it does not follow that the less important predictors of class have no role. They no doubt have the effect of reinforcing the main rules and so of contributing to the stability of the particular system' (1991:64).

Finally, in the limited instances where conflicts between formal and semantic criteria do arise, class systems normally resolve them in favor of the semantic criteria, giving meaning priority over form in assignment. Evidence for semantic priority comes from nouns whose meanings make them eligible for assignment by a semantic rule, but which have the 'wrong' form, i.e., nouns of a form not systematically associated with their semantic class. The classic example is Afar, an East Cushitic language described by Parker and Hayward 1985 and analyzed by Corbett 1991. Afar has a two-class system in which male humans and sex-differentiable animals are assigned to class A (*barisèyna* 'male teacher', *kùta* 'dog'), and female humans and sex-differentiable animals are assigned to class B (*baxà* 'daughter', *kuà* 'bitch'). The Afar classes also have formal correlates that overlap substantially with the semantic rules: nouns that end in a consonant or that are accented non-finally belong to class A (*gilàl* 'winter', *tàmu* 'taste'); nouns ending in an accented vowel belong to class B (*balillà* 'car', *gaqambò* 'bread'). In the few cases where the formal and semantic rules conflict, the semantic rules take priority. Thus, the noun *abbà* 'father' is assigned to class A even though it is accented on the final vowel like most class B nouns, and the noun *gabbixèra* 'slender-waisted female' is assigned to

¹²Arapesh* is Corbett's 1991:63 sketch of a 'non-attested system of conflicting assignment rules' made concrete with Arapesh particulars.

class B despite its non-final accent. Arapesh class assignment works the same way. In the rare exception to the generalization that semantic and formal rules overlap, the semantic rule takes precedence. So the Arapesh noun *nakor* ‘one’s own parent in-law’ is not assigned to class X as an *r*-final noun; nor does it receive default class VIII, as do most other nouns that are morphologically unusual.¹³ Instead, *nakor* receives class IV or class VII agreement depending on its intended referent, even though it does not meet the formal criteria for assignment to either of those classes (21).

- (21) *kwa-ti'* *nakor* *anan* *na-kih*
 IV.SG-see in-law VII.SG VII.SG-go.up
 ‘She saw her father-in-law arrive’

This is a fundamental continuity Arapesh shows with other types of classification systems.¹⁴

Given that formal assignment rules nearly always share their nominal territory to some extent with semantic ones, it is natural that languages would have some means for deciding which rule type will take precedence should they conflict. But why should they so consistently decide the same way? Is there is some general principle that makes semantic precedence the logical choice? One obvious candidate for such a principle would be the Elsewhere condition that governs interactions among sound-structural rules (see Anderson 1986; Kiparsky 1973; Koutsoudas, Sanders and Knoll 1974; and the many references cited in Janda 1987). The Elsewhere condition states that given two conflicting rules, Rule A and Rule B, if the environment of Rule B is a proper subset of the environment of Rule A, then Rule A will apply and Rule B will not. The application of Rule A will thus precede and disjoin the application of Rule B, and Rule B will apply in the remainder of environments, i.e., elsewhere.

The suggestion that class assignment might be governed by the Elsewhere condition was originally made by Aronoff 1994 in his analysis of the Lower Sepik language Yimas (Foley 1986, 1991). Like Arapesh, Yimas has an extensive set of rules classifying nouns on the basis of their final phonological segments, in addition to a core set of semantic rules, some of which also have systematic formal correlates. When formal and semantic criteria

¹³This example from Fortune’s texts refers to the husband’s father in ghost form. This does not bear on the noun’s classification, as the spirits of dead relatives are full participants in the Arapesh social world (see also Nekitel 1998b).

¹⁴Neset 2006:1385-86 hypothesizes that semantic priority in noun class assignment only holds for semantic distinctions that, as here, make reference to biological sex. He calls this narrower generalization ‘Core Semantic Override’.

conflict in Yimas, the semantic rule is what determines classification. For example, the Yimas noun *apwi* 'father' is assigned to masculine person class I on the basis of its meaning, rather than to the class specifically containing nouns ending in the segments *i* or *y*, Yimas class VIII. (Nor are nouns like *apwi* assigned to the language's largest and most formally diverse 'default' class V, which has been taking over ground traditionally held by class VIII.) Aronoff concludes that in Yimas the 'rules characteriz[ing] the assignment of nouns to the various phonologically based classes... all... stand in a default relation to conceptual [i.e. semantic] classes' (1994:116), offering the following by way of explanation: '[G]iven a noun with a conceptual property P_c and a phonological property P_p , why should the first take precedence? The answer is that nouns with both properties are a subset of those with only P_p , and so the conceptual definition takes precedence' (1994:189 n.30). If correct, this would be a significant finding, since Yimas, Arapesh, and virtually all other class systems give precedence to semantic assignment rules in the same way. Aronoff's suggestion implies that this pattern arises for reasons having to do with the logic of rule application, as opposed to being due to the nature of the specific linguistic categories involved.

Of course, the precise formulation of the rules in question is crucial for assessing the conditions that govern their application. A rule that preempts another according to the Elsewhere condition must apply to a proper subset of the cases to which the disjoined rule applies; otherwise the rules do not necessarily conflict. Any lack of overlap among the rules' environments is thus sufficient to exclude the Elsewhere condition from responsibility for the observed interaction, and another explanation must be appealed to instead. So to review the generalizations for Arapesh: (a) nouns denoting male humans are assigned to class VII; (b) nouns denoting female humans are assigned to class IV; (c) nouns denoting female humans regularly end in the phonological segment k^w (but not necessarily vice versa); and (d) nouns ending in the phonological segment n are assigned either to class VI or class VII depending on their semantics.

Semantic precedence is exhibited when conflicting assignments are predicted for the same noun like *məgan* 'sister-in-law' or *nakor* '(male) in law'. The male person semantics of *nakor* implies assignment to class VII, whereas its *r*-final form implies assignment to class X. In fact it is assigned to class VII, following the semantic rule. Similar cases can be cited from other Arapesh varieties; in other words, this is a common kind of exception. In Womsis Abu', *ufu'al* 'male spirit' and *baah* 'grandfather/grandson' normally receive agreement according to their male person reference, rather than according to their *l*-final or *h*-final form (22).

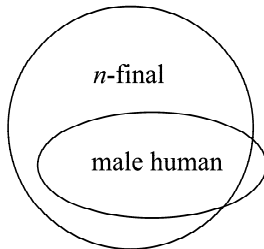
- (22) *ufu'al* *n-ahe'* (**l-ahe'*)
 male spirit M.SG-went

It is also common for both male and female human nouns borrowed from Tok Pisin to exhibit this pattern of semantic precedence when they have a final form not sanctioned by native Arapesh morphology (23).

- | | | | |
|-----------|-----------------------------|--------------------------------|-------------------------------------|
| (23) Abu' | <i>pater</i>
priest | <i>afu-neri</i>
good-M.SG | <i>n-ahe'</i>
M.SG-went |
| | <i>nes</i>
nurse | <i>afu-'i</i>
good-F.SG | <i>kw-ahe'</i>
F.SG-went |
| Mufian | <i>in-ini</i>
indef-M.SG | <i>misinari</i>
missionary | <i>n-angaf'o</i>
M.SG-went.there |
| Bukiyip | <i>en-ok</i>
indef-F.SG | <i>misis</i>
European woman | <i>k-onaki</i>
F.SG-came |
| Cemaun | <i>bisop</i>
bishop | <i>n-anaki</i>
M.SG-came | |

But is this pattern due to the fact that nouns with conceptual property P_c are a proper subset of those with phonological property P_p ? It seems not, because even when we limit our consideration to human nouns, those with male reference are not properly contained in the set of those with n -final form. Indeed, it is precisely the existence of exceptional nouns like *nakor*, represented by the segment of oval overextending the circle in (24), that provide the evidence for semantic precedence. Any time we find such examples we simultaneously lose the possibility of an explanation for class assignment in the regular cases in terms of the Elsewhere condition, since they show that the relationship between the two types of assignment rule is not one of proper inclusion.

- (24)



The idea behind the Elsewhere condition, self-preservation, is really a point of logic: if a rule or constraint that conflicts with another in whose environment it is properly included is not given precedence of application, it would never have a detectable effect on any form and could play no role in the grammar. The tendency for semantic class-assignment rules to take precedence over formal ones in situations of conflict must therefore somehow derive instead from the nature of the linguistic categories involved.¹⁵

What we are left with, then, is an irreducibly linguistic generalization about the distribution of class-assignment principles:

(25) Semantic assignment >> Formal assignment

The hierarchy in (25) expresses the notion that cross-linguistically, noun classes have a semantic core: if a language has only one non-default class assignment rule, it will refer to a semantic feature or a set of such features.¹⁶ There is no reason to expect that only one type of rule should apply for a given lexical item; indeed, conjunctive application is surely preferable, all else being equal, since it maximizes the applicability of rules of both types. But if a class system incorporates both types of rules, the semantic rules will generally have priority of application.

Why class systems should adhere to such a condition remains something of a mystery, though we can at least connect it to a body of work that seems to provide a constructive way of framing the issue. The hierarchy expressed in (25) can be understood as a specialized region of the referential hierarchy (Silverstein 1976, 1986, 1987; Aikhenvald and Dixon 1998). According to this view, all grammatical categories are located in a multi-dimensional space defined by cross-cutting arrays of linguistically relevant features. Within such a framework, 'overt systems of semantic CLASSIFIERS or covert systems of semantic NOUN CLASSES... are interestingly structured local subspaces in some particular region of [Noun Phrase-defining] categories' (Silverstein 1987:143). This point of view emphasizes that class is definable only in simultaneously formal and functional terms, so that formal classes of nouns that achieve the status of agreement categories must be anchored to some extent in meaning. In the absence of such anchoring, the generalizations embodied in a classification system are doomed to syntactic

¹⁵In cases where the Elsewhere principle really appears to be violated with pairs of productive morphological rules, the disjunctive ordering relations similarly call for some other explanation, like stipulation on a language particular basis.

¹⁶Examples of one-rule languages are found in the Kolami-Parji sub-group of Dravidian. Several of these (Kolami, Ollari, Parji) have one rule referring to male humans, and another class encompassing everything else; see Corbett 1991:10-11.

inertness; they lack the power to extend beyond the lexemes in which they inhere. Semantically unanchored formal generalizations can define lexical, but not morphosyntactic, classes.

But an anchor in meaning need not extend to every last category in the system, as Arapesh so clearly shows. Like the individual classes discussed above, the overarching category of noun class or gender has a structure in its own right, which we can see taking shape in the generalization in (25). As with other categories such as person and number, class is complex, anchored in features that are arranged hierarchically. Semantics has a place at the top of this hierarchy, playing a role in every class system.¹⁷ But once meaning is brought into play somewhere, secondary categories based solely in nominal form also emerge as a typological possibility. Indeed, as I will discuss further in the following two chapters, the number of categories rooted exclusively in form within a given class system is in principle unbounded, which is to say that the limits are set only by the language's ultimate formal means, its inventory of sounds.

Conversely, there seems to be a correlation between purely semantic systems and limited agreement. As Audring 2009 shows, gender systems that are supported only by pronouns are almost always purely semantic, suggesting that agreement is not just a reflection of grammatical class but also serves as a 'scaffolding device' allowing for the extension of form-based categories to other parts of the lexicon. The relative priority or power of semantic assignment rules is also seen in the Arapesh phenomenon of 'lexical splitting', whereby irregular and hence listed inflectional forms corresponding to a single lexeme proliferate in the system. We find massive lexical listing of kin term plurals in all Mountain Arapesh varieties, very frequent listing of singular kin terms due to the need to specify irregular direct vs. indirect (own vs. other's) forms for the same relationship, and in the southwestern Arapesh languages Weri and Balif Mufian, where all kinship nouns are inherently marked for their possessor, formal irregularities necessitate a listing of terms corresponding to virtually every cell in every noun's person/number/gender paradigm (Dobrin 1997). Why does this occur specifically in the kinship vocabulary? Nearly all native Arapesh nouns referring to humans also refer to kinship relations, making this the primary area of the lexicon where semantic class assignment rules hold sway. The semantic classification of nouns on the basis of the human status and sex of their referents is the morphosyntactic 'glue' that binds these terms together, helping constitute the lexicon as a structure, and not just a

¹⁷The range of semantic features that are potentially relevant to class is discussed in Aikhenvald 2000; see Adams and Conklin 1973 and Silverstein 1986 for proposals about their hierarchical organization.

list. At the same time, semantic classification allows the human nouns a degree of freedom that their nonhuman counterparts do not have: because it is only their semantics that affect their classification, the kinship nouns have fewer system pressures upon their form, so that lexical splitting can occur without disrupting that area of the lexicon as a whole.

Before leaving this topic, I should note that conflicts between rules of distinct types are not the only kind to arise in class assignment, and although giving semantics precedence is the solution languages most often adopt, what guides resolution can be rather less clear. For example the German noun *Mädchen* 'girl' is assigned neuter class presumably because of its diminutive morphology, and in spite of the fact that it denotes a female and should thus be feminine according to the relevant semantic criterion operative in German. The assignment of *Mädchen* is noteworthy in several respects. For one thing, it is not categorically neuter. *Mädchen* is a classic example of a hybrid noun, one that is not unambiguously assigned a unique class. Although *Mädchen* always takes a neuter form of the article, the noun can be treated as either neuter or feminine in pronominal agreement (i.e. selecting either *es* 'it' or *sie* 'she'). In a sense, then, the conflict among criteria is resolved not by following one rule at the total expense of the other, but rather with a compromise. This shows still another aspect of noun classification's semantic core: where formal assignment appears to take precedence over semantics, its ability to do so is 'weak'; it does not extend equally across the entire paradigm of agreeing forms.

Mädchen would seem to provide a prima facie counterexample to the generalization that semantics takes precedence over form in cases of conflict, because the noun is assigned to the class expected on the basis of its suffix, rather than on the basis of its meaning. But it is not the suffix's formal features *per se* that guide the class assignment of *Mädchen*.¹⁸ The suffix *-chen* also carries with it a diminutive meaning, so that the factor overriding the relevance of the noun's female denotation in class assignment could be either the suffix as such or the semantics it contributes. Evidence favoring the latter possibility comes from the fact that there are at least three formally distinct diminutive suffixes in standard German, *-chen*, *-lein*, and *-le*, all of which predict neuter class (*der Tisch* ~ *das Tischchen* 'table', *der Brunnen* ~ *das Brunnlein* 'fountain, spring', *der Garten* ~ *das Gärtle* 'garden'; see Drosdowski 1984:202, Hermann 1985:40). What these suffixes share is semantic, not formal, and their common influence on class assignment presumably derives from what they share. If that is the case,

¹⁸In the case of *Mädchen* there no longer exists an independent stem to which the diminutive suffix is attached. But the diminutive pattern in question is itself productive, as seen from pairs like *die Madam* ~ *das Madamchen*.

then we are dealing not with a conflict between semantics and form, but with a conflict between two different semantic categories.¹⁹

In Arapesh there are several nouns that approach but fall short of the threshold for semantic class assignment. One of these is *tutukwir* ~ *tutukwigu^h*, which denotes a specifically male ogre yet receives class X agreement following its form (*tutukwir* is also the name of a tag-type game played by children; with that meaning the noun's classification is unremarkable). It is not that ogres do not count as human, because *babamek^u*, a female ogre, is assigned to class IV as expected. Another below-threshold case is the set of nouns referring to very small children, which are built on formally appropriate class-marked stems: *aramatokwiñ* ~ *aramagowef* 'female infant', *aramaniñ* ~ *aramumwif* 'male infant'. These take agreement with default class VIII rather than with the class predicted by the natural sex of the referent. Evidently it would be more accurate to say that the semantic rules specifically register functionally male and female persons. Such a principle may also go some way toward explaining cases like Greek *agori* 'boy' and *koritsi* 'girl', which are neuter despite having an identifiable natural gender, apparently reflecting a semantic quality of relative smallness or immaturity if not morphological diminutivity (see Holton, Mackridge, and Philippaki-Warbuton 1997:248-249).²⁰

Within the set of semantic class assignment rules, logical considerations might be thought to exert an influence on rule interaction, though that possibility collapses under scrutiny as well. In Swahili, nouns denoting animates are assigned to class 1/2, but there are specific rules for augmentatives (+AUG → 5/6) and diminutives (+DIM → 7/8) that must take priority, 'otherwise nouns like the augmentative *j-oka* 'giant snake' and the diminutive *ki-toto* 'baby' would be assigned to class 1/2 by the [general] rule' (Corbett 1991:47-48). If the augmentative and diminutive rules didn't win,

¹⁹There are some additional arguments for dissociating the neuter class of diminutives from the suffixes themselves. The modern noun *die Küche* 'kitchen' ended in the sequence *en* as late as early New High German (Paul 1968[1917]) and thus had a phonetic form that was indistinguishable from a diminutive (especially given its unlauded stem). Nevertheless, the noun's class remained unaffected. According to Köpcke 1988:327-328, the subsequent loss of the final *n* resulted from a back-formation that occurred when *Küchen* was reinterpreted as a plural. The noun *die Kirche* 'church', which developed an *n*-final singular on analogy to its oblique case forms, later dropped the *n* again, apparently because speakers were avoiding its interpretation as a plural. Note again the strategy not adopted: *Kirchen* was not assigned neuter gender, an alternative means of affirming the noun's singular number that might have been employed if its *chen*-final form were exerting any influence on its interpretation.

²⁰Note that we are discussing here lexical class assignment, as distinct from the phenomenon of 'pragmatic projection' whereby a noun's lexical specification is overridden by a value associated with the lexical field as a whole (Zubin and Köpcke 2009).

when would they ever get to apply? But there are other, non-animate nouns that can be marked as increased or diminished in size, number or character, such as *j-umba* ‘large house’ (cf. *ny-umba* ‘house’) or *ki-jiwe* ‘pebble’ (cf. *ji-we* ‘stone’). Because the overlap between diminutive and augmentative assignment on the one hand, and general animate assignment on the other is incomplete, it is necessary to stipulate the priority of the former. In Swahili this can be accomplished by treating the diminutive/augmentative rules as class-changing rather than as applying to a raw stem. Where the prefix is not class-changing but rather serves to create a stem, like *ki-pofu* ‘blind person’ or *ki-wete* ‘lame person’, the rules are not in competition and the general animate rule holds sway.

A similar issue arises within a set of formal assignment rules. Do the *bør*-final nouns of Arapesh class II form a proper subset of the *r*-final nouns of class X? If class assignment only took account of singular form, the answer would seem to be yes; however, the two sets of nouns also have distinct plurals which are relevant to their categorization, undermining the totality of overlap in their environments (and removing hyper-inflected plurals like *teleboguhwas* discussed in note 2 above from the domain of possible Elsewhere violations).²¹ The Elsewhere condition thus appears not to play a role in organizing either of the two main types of Arapesh class-assignment rules.

Noun classes exhaustively partition the nouns in a language into categories. But what happens when a noun has none of the criterial properties referred to by the class-assignment rules and therefore does not fit into any category? A last-resort rule gathering up such nouns is a morphological default rule like the Arapesh rule providing class VIII agreement for nouns that are unassignable according to the regular rules. As we will see in the next chapter, there is yet another, qualitatively different type of default rule at work in the class assignment system of Arapesh. This rule is cross-linguistically unusual in that it refers directly to a noun’s concrete phonological form.

²¹ Recall that there is also a set of exceptional nouns that take *-ehas* in the plural and class VIII agreement, such as *burehør* ‘shell money’, *sarar* ‘grave platform’, and *tatuwar* ‘arrow’.

Arapesh Defaults

5.0 Default Categories

In this chapter, we consider the behavior of defaults in Arapesh noun classification. Morphological systems often have some ‘last resort’ mechanism that assigns a morphological mark or value (in the cases at issue here, a noun plural or class feature) to lexemes that do not meet any of the system’s regular assignment criteria. Each of the Arapesh varieties has such a last resort or default category that satisfies the syntactic requirement for agreement with nouns that would otherwise be left classless by the set of regular rules.

In Section 5.1, the plural suffix *-ehas* is discussed in further detail, focusing particularly on the question of its ability to apply without regard to the form of its base. As we will see, an adequate analysis of this morpheme suggest that lexical relations or schemas really stand behind even the most regular Arapesh plural marking pattern. To that extent the *-ehas* plural bears on the debate over how the differences between regular and irregular morphology should be explained.

Section 5.2 examines the notion of default as it plays a role in the expansion of the noun class systems of Mufian, Bukiyip, and Cemaun Arapesh. All these languages have countenanced a new singular agreement pattern that incorporates *s*-final borrowings and takes alliterative agreement with *s*. This innovation is interesting because there exists a native default agreement pattern, associated with class VIII, to which we might expect such borrowings to be assigned. The pattern in question suggests that there are actually two distinct types of form-based class-assignment rules coexist-

ing as solutions to the same problem in the Arapesh system. One assigns class to morphosyntactically uncategorized nouns by relying on the language's morphological resources; the other, more unusual type relies on phonology directly.

Finally, further pursuing the contrast established in Chapter 1 between the phonological basis of class assignment in Arapesh and its morphological basis in Bantu, Section 5.3 considers developments that have taken place in a well-described set of highly elaborated Bantu noun class systems (KiTuba, Kinshasa LiNgala, and the Shaba dialect of Swahili) in situations of contact. Though these Bantu noun classification systems show evidence of restructuring in both the morphological realization of nouns and their patterns of agreement, in no case do we find anything at all resembling the innovation of new agreement classes in accord with a phonological principle. That is to say, although these languages exhibit alliterative agreement patterns, they are superficial: they give no evidence that a phonological agreement principle is operating productively in the Bantu languages as it is in Arapesh.

5.1 More on the Default Plural *-ehas*

5.1.1 One Rule or Many?

What does it mean to say that class VIII is 'the default noun class' in Arapesh? When nouns take the class VIII suffix *-ehas*, the plural marker that appears on nouns not accounted for by the regular form-based rules, class VIII agreement is typically found as well. Tok Pisin borrowings of this type are shown in (1), from Cemaun, where they take the cognate plural *-əhəs* (see also the native Rohwim examples in Chapter 2 (19)).

- (1) *beŋ* ~ *beŋəhəs* 'bank'
 laŋ ~ *laŋəhəs* 'taro pounder'
 kek ~ *kekəhəs* 'cake'

The plural form that appears in such cases is presumably supplied by a default rule: 'Since their singular forms do not fall under the scope of any [plural realization]-class-assigning rules... they will automatically fall into the default class if every noun must bear [a realization]-class feature' (Aronoff 1994:106). As argued in Chapter 2, what makes the rule assigning plural *-ehas* special is that it is the least constrained of all Arapesh plural rules in the phonological conditions on its environment.

As discussed earlier, Janda 1990:142 uses the criterion of 'distributional breadth'—the ability of an element to appear in a range of diverse con-

texts—to support the claim that German plural *-s* ‘represents the unmarked, general-rule default case’, predicting on that basis that *-s* ‘will eventually spread to numerous other nouns’ in the language. But there are other qualities associated with unmarkedness that are not so clearly reflected in the application of the *-ehas* ‘default’ rule. One of those qualities is productivity: *-ehas* is not normally extended to new nouns when a more specific plural—a local default—could apply (i.e., the plural of the borrowed noun *sipun* ‘spoon’ is *sipunab* in Rohwim following the specific pattern for pluralizing *n*-final non-human nouns, and not **sipunehas*). In effect, an Arapesh plural is only marked in relation to a particular phonological context. While the rule adding *-ehas* may have an environment that allows it to apply anywhere, it is only because the more specific, context sensitive plural rules are assumed to take priority.

But *-ehas* must also be specified in the lexical entries of a formally diverse group of exceptional nouns that we would expect to be pluralized according to the regular form-based rules (and as we will soon see, these ‘exceptions’ can be quite systematic). An example from Cemaun is the noun *mour* ‘work, task’, whose expected plural is **mouguh^w*, but which surfaces instead as *mourəhəs* (this is comparable to the lexeme *sam* ‘taro and coconut croquette’). In this case, *-ehas* must be present in the lexical entry of *mour* to block the regular form-sensitive rule from applying. When *-ehas* occurs on otherwise normal-sounding nouns like Cemaun *mour* and Rowhim *sam* it appears as if the default has over-applied: evidently the diversity of environments in which *-ehas* occurs has led speakers to extend it into new domains, transforming it from a ‘savior’ plural that applies only when needed into a ‘usurper’ or ‘imperialistic’ (Gerhardt 1994) plural that actively competes with and blocks other locally regular forms. In the terms used by Fraser and Corbett 1997, the same marker functions both as a ‘normal case’ default filling in information that is absent from a lexical entry, and a listed element that is idiosyncratically and overtly specified (see also Booij 1999).

Does this mean that we need to posit more than one *-ehas* plural? The question is worth asking because although it seems to be descriptively necessary, the morphological rule supplying *-ehas* also constitutes the form/meaning association we are trying to unify in the notion of the morpheme. It is also the source of the unified morphosyntactic properties (class VIII agreement) associated with this form. We are left with the question of how the invariant properties of a single element—i.e., the fact that *-ehas* always implies the same class VIII value, regardless of how it is derived—can be expressed in the grammar. Once again, we have reason to treat these forms as listed or derived as their predictability requires, and express what they have in common by a redundancy rule of the form $X \leftrightarrow Xehas_{[+PL]}$. It

is this redundancy rule to which agreement makes reference, if we are right in following the familiar linguistic reasoning that treats elements that function together as sharing some unitary representation.¹

There is further evidence that the *-ehas* plural must be understood as a family of lexical patterns, some more rule-like and some more list-like, that is summarized by a redundancy rule or schema rather than composed by a structure-building rule. This evidence is worth considering insofar as it bears on the principles distinguishing regular from irregular morphology.

5.1.2 Is the *-ehas* Plural ‘Regular’?

It has been argued (Pinker and Prince 1994; Marcus, Brinkmann, Clahsen, Wiese, and Pinker 1995; Clahsen 1999; Pinker 1999; Pinker and Ullman 2002) that regular and irregular morphological patterns are computed by distinct mechanisms, the first making use of symbolic rules to concatenate a variable with an affix, the second consisting of representational overlap in a network of listed forms. This has come to be known as the ‘dual route’ or ‘words and rules’ model. An opposing view argues that regular and irregular morphological patterns are processed uniformly by an associative network that maps a phonological representation of the base onto a phonological representation of the inflected form (Bybee 1995; Plunkett and Nakisa 1997; McClelland and Patterson 2002; Nakisa, Plunkett, and Hahn 2002, Bybee and McClelland 2005). Arapesh plural marking is interesting from the perspective of this debate, since the system incorporates form-dependent patterns as the regular state of affairs, making only exceptional use of symbol-concatenating rules. In the terms established in this literature, Arapesh morphology can be characterized as having a ‘minority default’ in which frequency and default status are decoupled. As we will see in what follows, loan-pluralization data from Cemaun suggests that even the most generalizable Arapesh plural is distinguished from the form-based plurals by degree, and not by qualitatively distinct mechanisms as the words and rules model would predict.

The key basis for drawing the distinction between ‘regular’ and ‘irregular’ morphological processes in this literature is their differential sensitivity to input form. Irregular rules make reference to formal similarities among

¹If we find that certain sets of features consistently behave as a unit with respect to certain types of rules of assimilation or resequencing, we have good reason to suppose that they constitute a unit in phonological representation, independently of the actual operation of the rules themselves. There is a useful analogy here to syntax: many of the most enduring results of syntactic analysis have been made possible by the recognition that word-groups functioning as single units with respect to syntactic rules form hierarchical constituents in phrase-structure analysis’ (Clements 1985:226).

their inputs, as in the English verbs *blow*, *grow*, *throw*, and *know*, which all take irregular *-ew* past tense forms. Regular rules, by contrast, are not so discriminating, and can apply instead to any morphosyntactically appropriate element, regardless of its phonological shape. The words and rules model separates the processing of regular and irregular morphological patterns into two separate systems. Irregular morphology of the *throw* ~ *threw*, *grow* ~ *grew* sort is claimed to be a function of associative memory, expressed theoretically by treating generalizations over irregular forms as emerging out of structural similarities among the items rather than by the extraction of their shared features. Regular morphological patterns, on the other hand, would seem to be derived by symbolic rules that concatenate an affix with a variable that stands for the base. Such rules presuppose no particular relationships among their inputs. Examples of symbolic rules are the assignment of English plural *-S* and past *-ed*, and as we have analyzed it so far here, the assignment of the plural *-has* in Arapesh. The distinction has nothing to do with the number of words participating in a given pattern or the ‘markedness’ of the pattern in the overall context of the language’s morphology. Instead, degree of (ir-)regularity depends upon the formal nature of a pattern. If reference to the form of a base determines or affects its inflection, then the pattern is considered irregular.

Proponents of the words and rules approach bring a number of sources of evidence to bear in support of this regular/irregular distinction. Sensitivity to frequency, for example, is a well-documented property of irregular morphology, not only in terms of a pattern’s distribution over the vocabulary, but also in the results of a host of experimental manipulations. The naturalness ratings given by subjects to a regularly inflected form has been found to correlate strongly with the naturalness of its stem, from which it would be transparently derived by rule, while the naturalness ratings given to an irregularly inflected form correlates not as much with stem naturalness as it does with the form’s own frequency of occurrence. There is also evidence that similarity among base forms is important for irregular but not regular morphology. While the likelihood of experimental subjects extending an irregular inflectional pattern to a new base depends on its degree of similarity to existing forms (e.g., English present ~ past *spling* ~ *splung* is more likely than *spiv* ~ *spuv*; see Bybee and Moder 1983), Prasada and Pinker 1993 found that regular English past-tense patterns are extended equally to peripheral and to prototypical verb bases. On the dual model, sensitivity to features of the base would not be expected if the base consists only of a variable.

One type of morphological pattern with particular importance for this debate is the ‘minority default’, in which the respective influences of frequency and regularity (in the sense of application irrespective of the form of

the base) are dissociated because the regular symbol-concatenating ‘default’ rule applies only in a minority of cases. Minority defaults like the Arapesh plural *-ehas* pose a challenge to network models because they are relatively infrequent, making generalization difficult. So unlike the regular English noun plurals and verb inflections, which apply to the preponderance of bases, the regular German *s*-plural and Arabic suffixed ‘sound’ plurals (*naazih* ~ *naazih-uun* ‘successful male’, *taʕriif* ~ *taʕriif-aat* ‘definition’) as opposed to templatic or ‘broken’ plurals (*ʕunquud* ~ *ʕanaqiid* ‘cluster’, *ʕandaliib* ~ *ʕanaadil* ‘nightingale’) have been claimed to be exceptional within their respective systems in that they apply to only a minority of forms (see also Haspelmath 1989 on plural formation in Hausa). How could a child or pattern associator learn to supply regular inflection for novel forms when the new form is more likely to look like an irregular than a regular? And how will a network based on sensitivity to similarities develop a unified category for inputs that have no formal features in common?

There is evidence from connectionist models that rule-less systems are a viable learning mechanism even for systems with minority defaults (Plunkett and Nakisa 1997; Nakisa, Plunkett, and Hahn 2002; Feldman 2005). Questions have been raised about the ‘minority’ status of the Arabic sound plural (Boudelaa and Gaskell 2002), but one interesting result to emerge from this literature is that while minority defaults do not have the same degree of constraint on their distribution that regularly inflected forms do, nor do they apply to an entirely heterogeneous set of phonological forms. For example, Plunkett and Nakisa 1997:33 found that for Arabic plurals, ‘membership in the sound plural is predictable from the phonological form of the singular’ and that ‘singulars which take the sound plural inflection are just as coherently clustered in phonological space as many of the broken plural classes’ (though the strength of this effect depends on the method of calculation; see Boudelaa and Gaskell 2002:334-335).

Moreover, uniform processing models emphasize that type frequency is an important factor determining productivity (Bybee 1988, 1995, 2001). In reviewing the case for a symbolic plural rule suffixing *-s* in German (Marcus et al. 1995), Bybee 1995, 2001 makes an observation for German that is similar to the one made by Plunkett for Arabic: although the type frequency of the *s*-plural is quite low, and although it does apply to a wide formal range of inputs, *-s* is nevertheless extended to borrowings and nonce forms according to their phonological shape. The likelihood that a new form will take the *s*-plural is increased if existing lexical items that the new form resemble take the *s*-plural themselves, as is the case with proper names and words ending in a full vowel. Thus, the frequency of *s* across a morphologically defined subsets of nouns still affects productivity. The

existence of such ‘islands of reliability’ (Albright and Hayes 2003) suggests that speakers are not simply applying a single undifferentiated rule.

Although it appears on a minority of forms, *-ehas* is the sole plural in the Arapesh system that can be represented by a rule concatenating a suffix with a variable. All the other plural rules, including the ‘local defaults’ discussed in Chapter 2, are sensitive to the form of the base. It should therefore be possible to make predictions about how these plurals will be treated by speakers that bears on their processing. And we do have some evidence about the extension of the Arapesh plural rules from the treatment of loan words by native speakers.² If the words and rules model is correct, we should not expect frequency-sensitivity in the extension of *-ehas*, in contrast to choices among form-based plurals, which should depend on the degree to which the loans resemble existing forms.

Code-switching between the vernacular and Tok Pisin is now ubiquitous in Arapesh, and nearly every occurring Arapesh utterance presents an occasion to see how Tok Pisin borrowings are incorporated into an otherwise Arapesh syntactic frame in which number is normally obligatorily indicated (plural number is not expressed morphologically in Tok Pisin, but rather by preceding the noun with the third-person plural pronoun *ol*). For the most part, borrowed nouns are pluralized according to the native form-sensitive patterns, as the Cemaun examples in (2) show.

(2)	<u>Class</u>	<u>Singular</u>	<u>Plural</u>	
	XI	<i>tiket</i>	<i>tiketog</i>	‘ticket’
	VI	<i>pen</i>	<i>penəb</i>	‘paint’
	III	<i>baig</i>	<i>baigəs</i>	‘bike’
	IV	<i>hauskuk</i>	<i>hauskukmɛb</i>	‘cooking house’
	X	<i>wisil</i>	<i>wisiloguh</i>	‘whistle’
	VIIIa	<i>sia</i>	<i>siahəs</i>	‘chair’
	VIIIb	<i>kek</i>	<i>kekəhəs</i>	‘cake’

As can be seen from VIIIa and VIIIb, the *ehas*-plural (realized *-əhəs* in Cemaun) applies as in the native vocabulary to nouns ending either in a vowel or in a sound that is beyond the scope of the form-based plural rules (here, it is suffixed to a noun ending in non-labialized *k*). The same pattern is found in Mufian and Bukiyip. In Womsis Abu’, liquid-final borrowings may also take an *-ihəs* plural that is cognate to Arapesh *-ehas*, e.g., *sukul-ihəs* ‘school-PL’.

²The evidence is not experimental. I found native speakers unwilling to produce or evaluate the plurals of nonce forms they interpreted as belonging to their language. Instead, they would claim not to know the word or question the knowledge of the person whom I had heard utter it.

But there is also a set of curious borrowed nouns that do not seem to follow the native morphology:

(3)	<u>Bukiyip</u>		<u>Mufian</u>	
	<i>pasin</i>	‘custom’	<i>kar</i>	‘car’
	<i>sugul</i>	‘school’	<i>skul</i>	‘school’
			<i>baibel</i>	‘bible’

As shown in (4), these nouns receive default class VIII agreement rather than the agreement form expected on the basis of their phonological form. Most of the words in question end in a liquid. None of these nouns is recorded in the plural in Bukiyip or Mufian, though on the basis of their agreement we assume they take *-ehas*.

(4)	Bukiyip	<i>ch-ilawali</i>		<i>kipai-ñi</i>		<i>pasin</i>
		VIII.PL-will.bring		another-VIII		custom
		‘they will bring another custom’				
	Mufian	<i>kar</i>	<i>inimba</i>	<i>so ’u-so ’u-ñei</i>		<i>Toyota</i>
		car	this.VIII	small-small-VIII		Toyota
		‘this very small car is a Toyota’				

In Cemaun, we also have examples showing how liquid-final borrowings are pluralized. While many are assimilated to class X along the lines shown in (2), another subset take *-əhəs*:

(5)	<u>Singular</u>	<u>Plural</u>	
	<i>wilwil</i>	<i>wilwiləhəs</i>	‘bicycle’
	<i>sawol</i>	<i>sawoləhəs</i>	‘shovel’
	<i>sel</i>	<i>seləhəs</i>	‘canvas’
	<i>tawol</i>	<i>tawoləhəs</i>	‘towel’
	<i>skul</i>	<i>skuləhəs</i>	‘school’

What might account for this idiosyncratic overapplication of *-ehas*? The answer is that it reflects a native Arapesh tendency to apply this plural to liquid-final bases, suggesting that the ‘regular’ default rule is not as insensitive to the form of its input as we might have assumed. As it turns out, the domain covered by this plural is not randomly populated in the native lexicon. Of the sixteen nouns listed as belonging to this subclass in Fortune’s grammar, half end in a liquid (e.g., *gwar* ~ *gwarehas* ‘wig of ringlets’). There is also a liquid-final subclass in Bukiyip that takes the *-ehas* plural and agreement with class VIII (*miyokul* ~ *miyokulhas* ‘fish net’, *ati-ñ kwal*

‘one year’ ~ *wolobai-chi kwalahas* ‘many years’). The pattern is robust as well in Cemaun, where examples are plentiful (6):

(6)	<u>Singular</u>	<u>Plural</u>	
	<i>nugur</i>	<i>nugurəhəs</i> ~ <i>nuguh</i> ^w	‘chin’
	<i>mour</i>	<i>mourəhəs</i>	‘work’, ‘task’
	<i>aisir</i>	<i>aisirəhəs</i>	‘type of tree’
	<i>sur</i>	<i>surəhəs</i>	‘bush camp’
	<i>cər</i>	<i>cərəhəs</i>	‘clay pot’
	<i>wabienjər</i>	<i>wabienjərəhəs</i>	‘type of vine’
	<i>worimar</i>	<i>worimarəhəs</i>	‘type of yam’
	<i>nopunigur</i>	<i>nopunigurəhəs</i>	‘type of lizard’

The overapplication of the *-əhəs* plural to liquid-final borrowings implies that the rule is sensitive to these native lexical patterns. This is consistent with the uniform processing hypothesis, but unexpected if the plural is supplied by a regular rule that applies only to a variable.

There are also some cases in which even the most regular plural fails to apply in default. The result is a subset of borrowed nouns that have no distinct plural form. Borrowed nouns ending in *s* systematically fail to receive any plural marking at all (7), as do many nouns that end in a round vowel (8).³

(7)	<u>Noun</u>	
	<i>bisnis</i>	‘business’, ‘work for money’
	<i>krismas</i>	‘Christmas’, ‘party’, ‘year’
	<i>tinpis</i>	‘canned fish’
	<i>katres</i>	‘bullet’
	<i>kes, sutkes</i>	‘suitcase’
	<i>dis</i>	‘dish’, ‘plate’, ‘wash basin’

(8)	<u>Noun</u>	
	<i>so</i>	‘saw’
	<i>su</i>	‘shoe’
	<i>pilo</i>	‘pillow’
	<i>hango</i>	‘Japanese saucepan’

This failure to inflect carries over to a few borrowed nouns of other forms, including *brum* ‘broom’, *sigaret* ‘cigarette’, and *sop* ‘soap’, though

³Not all such nouns fail to pluralize; cf. *pato* ~ *patohəs* ‘duck’, *bairo* ~ *bairohəs* ‘pen’.

the pattern is less consistent (cf. *rum* ~ ?*ruməhəs* ‘room’, *mit* ~ *mitog* ‘meat’, and *kap* ~ *kas/kasəhəs* ‘cup’). The status of such nouns as count nouns seems undeniable when they are introduced with a numeral or other quantifier: *ət-u su* ‘one-AGR shoe/pair.of.shoes’, *bie-s tinpis* ‘two-AGR can.of.fish’; *worubai-wi su* ‘many-AGR shoe/pair.of.shoes’. In the next section we will explain the phonological agreement patterns of *s*-final borrowings by appealing to the fact that they do not participate in a number alternation, so they default out of the morphological class assignment system altogether, taking phonological agreement instead.

But the reason for the failure of *-ehas* to apply in the first place remains to be explored. If the rule is really insensitive to the form of the base, it should apply blindly whenever no other plural can. So then why the hesitation to apply to *s*-final and *u*-final nouns? The answer, it appears, is that because of their final segments these nouns already sound plural; i.e., they each conform to a formal plural schema in the language, and are therefore felt not to require further plural inflection. The pattern is very similar to the one described by Köpcke 1988:325 for loan words that take the zero-plural in German:

[Z]ero surprisingly appears to be overgeneralized, since this morpheme is not possible for monosyllabic nouns in the native lexicon. *Drops* ‘gum drops’, *Keks* ‘cake [cookie]’, and *Quiz*, which take the zero plural in alternation with a plural suffix, all match the plural schema based on *-s* in their singular form, a situation in which... experimental subjects frequently used zero in [a] nonce word experiment.

Apparently ‘speakers perceive the stem-final *-s* as a plural “marker”..., which creates problems for a theory of morphological segmentation, but apparently not for speakers of German’ (Köpcke 1988:325 n.12). It should cause no problem for speakers of Arapesh either, since their language keys morphological values to stem-final segments as the ordinary state of affairs. The failure of *-ehas* to apply to *s*- and *u*-final borrowings is an argument that in addition to being ‘input-oriented’ in having form-sensitive rules, Arapesh noun inflection is simultaneously ‘output-oriented’ in utilizing plural schemas that exist independently of their corresponding bases.

The conclusion that Arapesh plurals should be understood not as productive rules but as schemas over lexically listed forms comes as no surprise, since we have already argued that schemas are needed to unify the unpredictable yet motivated plurals that correspond to both phonologically specified categories of singular form and morphosyntactic classes. What is surprising is that they are able to block pluralization with *-ehas*, showing that even the system’s default plural does not have a completely open structural description. While at least some *-ehas* plurals would have to be listed to block the form-sensitive plurals from applying, the proportion of ‘listed

defaults' is relatively high, given the application of *-ehas* to a substantial liquid-final class. Moreover, the *-ehas* plural is not evenly distributed across singular form types. Rather than applying like a default rule just where it is needed, the *-ehas* plural is a family or cluster of summary descriptions that reference formal features of the base; it subsumes at least the two main form-dependent patterns in (9):

- (9) /XV/ ↔ /XVhas/[+PL]
 /XI/ ↔ /XIlehas/[+PL]

In other words, what makes the *-ehas* plural special with respect to the rest of the system is the fact that, unlike the other plurals, it fails to correspond to a phonologically unified class of singular forms, and it overlaps in its formal territory with the other form-based rules. It is still the plural with the greatest flexibility in terms of what it applies to, which is why it often functions like a default. But that does not mean that it must be regular in the sense of being totally open in its structural description. And indeed, in at least certain cases we know that it is not, as when it applies to nouns whose form should make them eligible for the more typical 'irregular' plurals. So the more we know about Arapesh plurals, the less we find evidence for structure-building rules, and the more it looks like lexical associations or schemas all the way down.

5.2 Default Agreement in Arapesh

The analysis of Arapesh presented in Chapter 2 takes the productivity of the different types of class assignment rules to be an important diagnostic for determining how the system is organized. Since form-based noun classification and plural-assignment rules are productive in the sense that they apply to new vocabulary, I have argued that it is really form-based classification that constitutes the normal situation in Arapesh, and only when class cannot be assigned on the basis of a noun's form does class VIII come to the rescue, assigning a value in default. In what follows I explore notion of form-based classification in more detail, leading to an enriched understanding of what can be meant by 'a noun's form' and 'default'.

5.2.1 Types of Formal Noun Class Assignment

Corbett 1991:51 provides the following heuristic to distinguish two types of gender or noun class assignment that are sensitive to a noun's form:

[I]f in order to establish the gender of a noun we need to refer to more than one form... then we are dealing with a morphological assignment

rule. If, on the other hand, gender can be established by reference to a singular form, then we are dealing with a phonological [assignment] rule.

That is, morphologically based rules require reference to a paradigm or to some morphological element the word contains, whereas rules that are based on phonology refer to an individual noun form as such.

The Arapesh noun class assignment rules are clearly morphological according to Corbett's criterion. Patterns of plural formation that are sensitive to singular-final phonological form are the basic structures to which the noun class categories refer. Nouns whose singular-plural relationship is unpredictable—not sanctioned by a rule—receive default classification. Nouns that take the plural *-ehas* are similarly classified by default, because the *-ehas* plural does not correspond to a phonologically coherent category of singular form. So to summarize, even though the classes correspond with high regularity to sets of singulars sharing same word-final phonological form, class assignment normally takes account of plural form as well: the Arapesh assignment rules are of the schematic form $\{/A/_{[+SG]} \leftrightarrow /B/_{[+PL]}\} \rightarrow$ class C, as shown below in (10). Within this set of rules, order of application can be regulated by degree of specificity, so that the ranking of morphological over default assignment displayed in (10) need not be stipulated. (Recalling the discussion in Chapter 4, it is not a logical necessity that semantics takes priority over formal assignment. Nevertheless, we find this to be a ubiquitous characteristic of noun classification systems.)

(10)

<u>Morphological Class Assignment</u>	>>	<u>Default Assignment</u>
$\{/Xb^v/_{[+SG]} \leftrightarrow /Xbys/_{[+PL]}\} \rightarrow$ I		$\{/X/_{[+SG]} \leftrightarrow /Xehas/_{[+PL]}\} \rightarrow$ VIII
$\{/Xb\varnothing r/_{[+SG]} \leftrightarrow /Xryb/_{[+PL]}\} \rightarrow$ II		
$\{/Xg/_{[+SG]} \leftrightarrow /Xs/_{[+PL]}\} \rightarrow$ III		
$\{/Xk^u/_{[+SG]} \leftrightarrow /XV_{[+LAB]}/_{[+PL]}\} \rightarrow$ IV		
$\{/Xm/_{[+SG]} \leftrightarrow /Xp^i/_{[+PL]}\} \rightarrow$ V		
$\{/Xn/_{[+SG]} \leftrightarrow /Xb/_{[+PL]}\} \rightarrow$ VI		
$\{/Xn/_{[+SG]} \leftrightarrow /Xm/_{[+PL]}\} \rightarrow$ VII		
$\{/X\tilde{n}/_{[+SG]} \leftrightarrow /Xj/_{[+PL]}\} \rightarrow$ VIII		
$\{/Xp^u/_{[+SG]} \leftrightarrow /Xs/_{[+PL]}\} \rightarrow$ IX		
$\{/Xr/_{[+SG]} \leftrightarrow /Xgu^h/_{[+PL]}\} \rightarrow$ X		
$\{/Xt/_{[+SG]} \leftrightarrow /Xg^u/_{[+PL]}\} \rightarrow$ XI		
$\{/Xuh/_{[+SG]} \leftrightarrow /Xruh/_{[+PL]}\} \rightarrow$ XII		
$\{/XVh/_{[+SG]} \leftrightarrow /XV_{[+FR]}/_{[+PL]}\} \rightarrow$ XIII		

What then of the generalization that Arapesh classes are rooted in noun-final phonological form? Before attempting to answer, we should pause to consider what the classes actually entail. Following Hockett 1958:231, Dixon 1982:160-165, Silverstein 1986:501, Corbett 1991, and many others, we have defined noun classification as a partitioning of the nouns in a language according to their patterns of syntactic agreement. While classes may have correlates in the form or meaning of nouns, the decisive criterion is always to be found outside the item itself, in the associated markers of agreement. But this definition underdetermines the Arapesh noun classification facts in one important respect: it says nothing about the substance of the agreement-markers; it only says that whatever form they take will be the same for all nouns of a given class.

But in Arapesh, it is not only a noun's agreement paradigm that is predictable by reference to its form; the concrete realization of agreeing elements is also normally predictable from the noun's form. That is, in a way similar to the noun class systems of many African languages, agreement in Arapesh is 'alliterative'. Examples of alliterative agreement from the Kru language Godié (also spelled Godie; see Marchese 1986, 1988) and the Bantu language Swahili (Welmers 1973) are presented in (11) and (12) below. Examples of alliterative agreement from the Arapesh legends published in Fortune 1942 follow in (13). Indicators of class on nouns are highlighted in bold; agreement markers are underlined. All nouns in the rest of this section are singular unless overtly indicated.

(11) Godié

<i>nyukpə</i>	<i>kəd-ə</i>	<i>nə</i>	<i>nii</i>	<i>mle</i>	<i>kəd-ε</i>
man	big	this	saw	animal	big

'the big man saw this big animal'

(12) Swahili

<i>ki-kapu</i>	<u><i>ki-kubwa</i></u>	<u><i>ki-moja</i></u>	<u><i>ki-lianguka</i></u>
basket	large	one	fell

'one large basket fell'

The rule in (14) is certainly odd. It is similar to the regular class-assignment rules of (10) in that it makes reference to a noun's final phonological element to predict agreement, but it does so without mediation through a morphosyntactic class feature: it predicts the form of agreeing elements directly from the final form of the noun. Furthermore, the rule looks only at a single noun form, irrespective of its morphological constitution or paradigmatic placement, and it does not care about the morphological status of the noun-final C—it is simply a segment of sound.

One might wonder whether the identity relationship this rule is meant to capture is really a linguistically significant generalization that ought to be expressed in the grammar, or whether it is simply an incidental synchronic reflection of historical developments like those that produced the English clitics *-s*, *-d*, and *-l* from the corresponding full form auxiliaries *is*, *had*, and *will*. The latter is what we might conclude in the Swahili case, for example, since although the pattern illustrated in (12) is alliterative, it is not carried throughout the system—i.e., several noun-class markers have zero allomorphs, and in many class paradigms the agreement markers are diverse, with some being phonologically identical to the noun marker and others not (for example, Swahili class 9 nouns, which are marked with *n-* or \emptyset , take pronominal agreement with *i-*, class 11 nouns, which are marked with *u-*, take adjective agreement with *m-*, and so on). Such divergences are presumably the reason why the formal identity of alliterative markers is not usually raised in the analysis of agreement, even though it is quite common to find phonological correspondences between person/number markers serving different functions, that is, the same forms serving as argument markers on verbs and in possessive constructions, or as markers of both subject and object (see, e.g., Siewierska 1998, 2004). Rarely do we find correspondences anywhere near as thorough as they are in Arapesh, where they not only carry through the entire paradigm of agreement forms for every class but also basically exhaust the language's phonological inventory. We would be hard put to conclude that Arapesh subject agreement prefixes derive historically from the corresponding independent pronouns—or the other way around. Synchronically, both simply fill a templatic slot in the marker with the same consonantal phoneme. The independent pronoun has the regular shape $\partial C \partial C$, with class indicated by the phonological segment instantiating the variable C; the subject prefix can be specified simply as C-. This is discussed in more detail below.

Assuming that a rule like (14) should be represented directly in Arapesh grammar, its role is obviously a limited one given the simultaneous existence of the morphological assignment rules in (10). Moreover, although alliterative agreement is the norm in the language, the system does tolerate some significant divergences from the alliterative ideal. These occur most

frequently in the plural, where due to the extensive plural allomorphy a single agreement marker may correspond to a range of noun endings, as shown in (15).

- (15) a. *pupwijameb* *wa*-*k^u*
 insects.IV.PL IV.PL-IV.SG
 ‘insects (attacked) her’
 (SG *pupijak^u* ~ PL *pupijameb* ‘insect’)
- b. *barahijer* *wa*-*kana* ‘*e babwen...*’
 granddaughter.IV.PL IV.PL-say ‘oh grandfather...’
 ‘the granddaughters said “hey grandfather”’
 (SG *barahok^u* ~ PL *barahijer* ‘granddaughter’)
- c. *unaryb* *wa*-*ki^h* *wa*-*rahbys-an*
 cassowaries.IV.PL IV.PL-come.up IV.PL-threaten-VII.SG
 ‘cassowaries came up and threatened him’
 (SG *unaruk^u* ~ PL *unarib* ‘cassowary’)

In each of these cases we find a constant form of the class IV plural subject prefix on the verb although the nouns themselves are marked with plurals of various forms. Nevertheless, I will argue that a rule like (14) really is active in the grammar of Arapesh, as it accounts for an interesting development that has taken place in several Arapesh agreement systems.⁴ This development was first documented by Robert Conrad and his colleagues (Alungum, Conrad, and Lukas 1978, Conrad 1987, Conrad and Wogiga 1991) based on Mufian and Bukiyip data collected during the 1970s, and it is confirmed by evidence from Cemaun Arapesh as spoken in the late 1990s. The phenomenon, which involves nouns borrowed from Tok Pisin, provides evidence for a default class assignment principle that is phonological according to the criterion laid out above, with the assignment of class to nouns and the realization of class on agreeing elements preceding hand in hand.

5.2.2 The Special Behavior of s-final Borrowings

Fortune’s documentation of Rohwim Arapesh derives from the early 1930s. At that time, ‘no great number’ of Tok Pisin nouns had been borrowed into the language (Fortune 1942:7); some of the borrowings documented from that period are presented in (16).

⁴There is a high degree of consistency across Arapesh varieties that motivates the comparisons here. Weri is the sole outlier, with formal classification completely dropped and all non-human nouns assigned to in a single class.

in (18). Bukiyip examples are from Conrad and Wogiga 1991 or Conrad 1987; Mufian examples are from Alungum, Conrad, and Lukas 1978.⁶

(18) Bukiyip

- a. *əgidak* *nebe-g-ali* *trag*
 this.III.SG big-III.SG-lasting.quality truck.III.SG
 ‘this big truck’
- b. *tagas* *kobi* *pu-wolu-gas*
 water.tank.III.PL NEG.IMP 2PL-IMP.break-III.PL
 ‘don’t you guys break the water tanks’
- c. *ya-bo* *əna-ti* *barit*
 1SG-hit INDEF-XI.SG ditch.XI.SG
 ‘I ran into a ditch’
- d. *bio-g^w* *betog^w*
 two-XI.PL bed.XI.PL
 ‘two beds’

Mufian⁷

- e. *wig* *naman-i-ngei*
 week.III.SG tomorrow-POSS-III.SG
 ‘next week’
- f. *wikifah* *angahemba* *bia-ngah*
 week.III.PL these.III.PL two-III.PL
 ‘these two weeks’

The semantic class assignment rules continue to take priority over rules based in form. Thus, for example, Bukiyip nouns like *kiap* ‘patrol officer’ and *wantok* ‘compatriot’ (19a,b) take class VII agreement irrespective of their form because they refer in these contexts to men.

⁶As elsewhere, the transcription of the original sources is generally retained. However, the Bukiyip schwa phoneme is written here as ə (rather than é), and the high central vowel is written here as i rather than as ú as in Conrad and Wogiga’s practice. Mufian and Bukiyip cognates are identified using the class-labeling scheme developed by Fortune 1942 for Rohwim Arapesh.

⁷Mufian voiced stops pre-nasalize intervocalically. The intervocalic *k* in *wikifah* is anomalous and may reflect standard Tok Pisin orthography (see Mihalic 1971).

(19) Bukiyip

- a. *mo-ne* *laigim-an-i* *kiap*
 1PL-*do* desire-VII.SG-REL patrol.officer.VII.SG
 ‘the patrol officer we wanted’
- b. *yek-i-ni* *wantok*
 1SG-POSS-VII.SG compatriot.VII.SG
 ‘my (male) compatriot’

Mufian⁸

- c. *in-in* *misinari* *n-angafɔ'o*
 INDEF-VII.SG missionary VII.SG-habitually.went.there
 ‘a missionary went there’

Class VIII continues to act as the default, absorbing most exceptions.⁹ For example, the noun *taia* ‘tire’ receives default class VIII agreement because it is vowel-final; the noun *laik* ‘choice, desire’ does as well because its final obstruent lacks the labialization which would make it eligible for assignment to class IV. So the basic structure of the system remains intact, applying to both native and borrowed nouns.

- (20) a. *taia* *əma-ñ*
 tire heavy-VIII.SG
 ‘a heavy tire’
- b. *apak-i-ñ* *laik*
 we-POSS-VIII.SG choice
 ‘our choice; our decision’

But there has been an innovation that can be seen most clearly in the classification of borrowed nouns ending in the phoneme *s*. The significance of this development lies in what it implies about the native Arapesh class system, elucidating the crucial role of phonological form both in defining and realizing the classes.

While *s* functions as a noun plural and plural agreement marker in all Arapesh varieties, singular-final *s* is rare. Accordingly, nouns ending in this sound are pluralized according to the default rule in the two cases that are attested in Fortune’s original data. One of these is the native noun *pas* ‘taro

⁸The absence of the high central vowel *i* at the end of the determiner distinguishes the class of this phrase from default marking. So *in-ini hetman* ‘INDEF-VII.SG leader’ contrasts in class with *in-in mitin* ‘INDEF-VIII.SG meeting’.

⁹Default agreement is marked by SG *ñ* ~ PL *ch* in Bukiyip, and by SG *n* ~ PL *s* in Mufian.

pounder', which takes the default plural *-ehas* (yielding *pasehas*) and follows the class VIII default agreement paradigm. The other is the borrowed noun *hors* 'horse' already mentioned in (16) above, which Fortune 1942:8 says assimilates to class VIII 'for all syntactical functions'.

In (21) we see a list of *s*-final nouns that are attested in more recent Arapesh data. All are unambiguously Tok Pisin borrowings. The *s*-final borrowings are treated in the same innovative fashion in both dialects.

(21)	<u>Bukiyip</u>		<u>Mufian</u>	
	<i>balus</i>	'airplane'	<i>kos</i>	'course'
	<i>rais</i>	'rice'	<i>bas</i>	'bus'
	<i>kes</i>	'box'	<i>tokples</i>	'vernacular language'
	<i>glas</i>	'windshield'		
	<i>opis</i>	'office'		

These *s*-final borrowings are not pluralized with the default suffix *-ehas*. Nor do forms agreeing with them reflect default assignment, as everything we have seen thus far would lead us to expect. Rather, borrowed *s*-final nouns exhibit no distinct plural form and take agreement forms with *s*, resulting in collocations such as those presented in (22), *ananis kes* 'his suitcase' and *balus sabih* 'the airplane landed'.¹⁰

(22)	<u>Bukiyip</u>					
a.	<i>anan-i-<u>s</u></i>		<i>kes</i>			
	VII.SG-POSS- <i>s</i>		suitcase			
	'his suitcase'					
b.	<i>yek</i>	<i>ye-ne</i>	<i>skelim</i>	<i>rais,</i>	<i>aligi</i>	<i>ye-yata-<u>s</u></i>
	1SG	1SG-did	distribute	rice	until	1SG-finished- <i>s</i>
	'I distributed all the rice'					
c.	<i>balus</i>	<i><u>sa</u>-bih</i>				
	airplane	<i>s</i> -came.down				
	'the airplane landed'					

¹⁰Gerstner 1963 documents one example of this pattern: TP *klos* (written *kloth* 'clothes, dress, European style clothing') receives agreement with *s*, e.g., *auzibal-isi kloth* '[blood]red-*s* dress'. (The dates of Gerstner's research overlap with Fortune's, though the two worked in different places; see Burgmann 1963.) Why *s*-agreement was found by Gerstner but not Fortune is not clear; it may be that the coastal village Gerstner visited had a longer history of contact with Tok Pisin. In any case, the attestation of *s*-agreement in the Arapesh of the 1930s supports the claim made below that it reflects a process already latent in Arapesh grammar.

- d. *nebebe-si* *balus* *sa-naki* *Ukarumpa*
 very.large-s airplane s-came.from Ukarumpa
 ‘a very large airplane arrived from Ukarumpa’
- e. *owiñ* *ahah-i-s* *opis*
 below over.there-POSS-s office
 ‘the office down below there’

Mufian

- f. *kos* *işimba*
 course this.s
 ‘this course’
- g. *bas* *sa-fi’i* *a-nda’* *pasim-as*
 bus s-came 1SG-did hold-s
 ‘the bus came and I stopped it’

There are only two exceptions to this pattern in the Bukiyip and Mufian texts. The noun *misis* ‘European woman’ refers to a female person so is assigned to class IV in conformity with the semantic rules (23a). And in one instance the noun *bisnis* receives agreement with class VIII, apparently due to semantic association with the noun *moul* ‘work’ (23b).¹¹

(23) Bukiyip

- a. *ən-ok* *misis* *ko-naki* *kwa-salik-e*
 INDEF-IV.SG white.woman IV.SG-came IV.SG-asked-1SG
 ‘some white woman came (and) asked me...’
- b. *ən-eñ* *bisnis*
 INDEF-VIII.SG work.for.money
 ‘some cash-cropping’

Since the sound *s* does occur as a native agreement marker (or part of a marker) with the plurals of three classes, I, III, and IX, could it be that these *s*-final borrowings are being interpreted as plurals? None of them have back-formed counterparts indicating that a noun like *bas* ‘bus’ is interpreted as the plural form of a lexeme whose singular form is **bag* (class III) or **bap* (class IX). To the contrary, the broader contexts of many of these examples confirm that the *s*-final borrowings are understood to be singular

¹¹As mentioned above, *moul* is a native noun that takes class VIII agreement. *Bisnis* and *moul* are attested together in Bukiyip: *mouli bisnis*, and *bisnisin moul* (‘business work’), both referring to cash cropping, i.e., work for money as opposed to subsistence.

in reference.¹² Yet *s* does not function as a native singular agreement marker in any dialect. Since Arapesh has morphological resources for class assignment which continue to function fully even in the modern languages with heavy borrowing, and these include a highly productive default class which is used with borrowings of most other sorts, why don't *s*-final borrowings receive ordinary class VIII agreement by default?

5.2.3 The Phonological Basis of the Arapesh Classes

The development of *s*-agreement for borrowed nouns ending in *s* is not actually new at all, but instead reflects the work of a native phonological agreement principle. Its role is to demand syntactic agreement not with a head noun's abstract class feature, i.e., with a morphological specification, but rather with its final consonant, i.e., with a phonological specification. The result is true alliterative concord: instead of separating the assignment of class in nouns and the realization of class in agreement into distinct processes or conditions, the former morpholexical, the latter morphosyntactic, the rule refers to the noun's phonological representation directly, making use of the same phonological features that are predictive of class in agreement. If this were the only agreement rule in Arapesh, then knowledge of a noun's phonological form is all that would ever be necessary to complete the specification of agreeing elements in their syntactic contexts. Since many agreeing elements contain some phonological material (especially vocalic material) in addition to the target noun-final segment, the lexical representations of these elements resemble partially specified templates of a sort familiar from reduplicative and root-and-pattern morphology. The agreeing elements listed in (24) in their partially specified lexical forms reflect the Rohwim dialect documented by Fortune.

(24)	Possessive	→	<i>-i-C</i>
	Intensive Pronoun	→	<i>aCaC</i>
	Verb Subject	→	<i>C-</i>
	Verb Object	→	<i>-C-, -C</i>
	Numeral	→	<i>-C</i>
	Adjective	→	<i>-Ci, -Cali, -Ceri</i>

¹²For example, the phrase including the *s*-final borrowing *balus* 'airplane' in (22c,d) appears with the following running gloss: 'And so we waited for the aeroplane. We waited and soon Wanguen and his children came... in a truck. They came. They came and we stood around. We stood around and soon the aeroplane came and landed. It came from Ukarumpa, which is far away. The white man, who was the pilot, brought some of Wanguen's things. It was a very large aeroplane and it came from Ukarumpa' (Conrad and Wogiga 1991:273).

Demonstrative

(near me)	→	<i>aCuda'</i>
(near you)	→	<i>naCuda'</i>
(near him/her/it)	→	<i>CaCuda'</i>

The consequences of positing a phonological agreement rule to fill the consonant slots in these templates will be explored in Chapter 6. In the meanwhile, we should further consider its justification given that the Arapesh classes are defined morphologically by way of abstract class features via the conventionalized pairings in (10). Recall, for example, the frequent incongruities between plural marking in agreement and plural marking on nouns, which are obviously not derived by way of rule (14). The problem is that without such a rule, we are unable to capture the phonologically exhaustive nature of the classification system.

The Arapesh consonant inventory in (25) is based on the description in Fortune (1942); the Bukiyip and Mufian inventories differ from it somewhat, though not in ways that affect the argument.¹³ The boxed-in segments in the chart represent gaps in the class system, i.e., consonants that do not predict a class. (The sound *k* is discussed separately below.)

(25)

<i>p</i>	<i>t</i>		<i>k</i>				
<i>b</i>	<table border="1" style="display: inline-table; vertical-align: middle;"><tr><td><i>d</i></td><td><i>j</i></td></tr><tr><td><i>s</i></td><td><i>f</i></td></tr></table>	<i>d</i>	<i>j</i>	<i>s</i>	<i>f</i>		<i>g</i>
<i>d</i>	<i>j</i>						
<i>s</i>	<i>f</i>						
<i>m</i>	<i>n</i>	<i>ñ</i>	<i>h</i>				
	<i>l-r</i>						

(*u* ~ *w*, *i* ~ *y* as a function of syllabic position)

The two boxes in the chart correspond to two straightforward generalizations. The sounds *d* and *j* do not appear word-finally on Arapesh nouns in either the singular or the plural; nor do they appear finally on words of any other lexical category. The fact that these sounds fail to participate in noun classification is thus due to their absence from the relevant class-defining position for phonotactic rather than morphological reasons. The segments *d* and *j* comprise a natural class within this system as the only two distinctive-

¹³Rohwim *f* corresponds to Bukiyip *c* and, in the native vocabulary, Mufian *s* (e.g., R *batauíf* ~ B *batíwíc* ~ M *batíwís* 'children'). Rohwim plural *s* corresponds to Mufian *h* (e.g., R *nugugwes* ~ M *nongwatoh* 'knives', R *bòbys* ~ M *bembéh* 'betel nuts', R *lawas* ~ M *lawah* 'tree'). The important point is that *s* does not occur word-finally on singular nouns in any Arapesh variety, even though it is phonologically admissible in this position in all of them.

ly voiced coronal obstruents, so that their absence from word-final position can thus be expressed by a specifically phonological feature co-occurrence restriction.¹⁴

The other box in the chart is the one containing *s* and *f*. These sounds appear word-finally quite commonly, but not on singular nouns. As a consequence, *s* and *f* serve as agreement markers only in the plural. With these distributional facts we can begin understand why speakers of Arapesh treat the new *s*-final nouns as they do. The gap in the class system corresponding to native *s* is a morphological accident, an anomalously missing category which could exist, given the phonological resources of the language, but which for some reason happens not to. With the introduction of a sufficient number of new nouns to warrant filling in that gap, the system has responded by sanctioning agreement with *s*, just as rule (14) predicts.

Evidence from Cemaun suggests that this phonological default may play a more central role in Arapesh class assignment/agreement than the material frozen in Fortune's older texts is able to indicate. There is a handful of native nouns that have only one form, sometimes but not always for semantic reasons (many are non-count, but this is not always the case). In such cases, agreement is regularly alliterative, as in (26a), even when a semantic number distinction can be sustained, as in (26b). In Cemaun the phonemes *g*, *k*, and *h* each have a labialized counterpart, *g^w*, *k^w*, and *h^w*. Of these six phonemes, *g*, *k^w*, *h* and *h^w* define singular noun classes; *g^w* defines a plural class. The sound *k* is the only consonant in the set that does not participate in the classification system. This seems to be tied to its unusual phonological properties, which extend across the family. As Fortune 1942:1 explains for Rohwim, 'the *k* phoneme is peculiarly unstable, and covers a wide range of sound'. In proximity to liquids it merges with *g*; in other environments it fricates to *ɣ* or to uvular *ʁ*. According to Fortune, at the time of his research *k* was completely omitted in speech used with infants. (Unfortunately we have no documentation of this special register and no opportunity to document it now given the language's advanced state of attrition.) In Cemaun *k* does not show this distribution, but it does have a tendency to delete in word-final position. There are prosodic, morphosyntactic, and lexical conditions on the deletion that are beyond the scope of the discussion here. The main point to note is that it is the only consonant in the Cemaun inventory that behaves in such a volatile way.

Phonological agreement is found with many borrowed nouns that can be interpreted as conforming to any of the native consonant phoneme cate-

¹⁴Cemaun has two nouns ending in the palatal alveolar affricate *j*, *kəmijawij* ~ *kəmijawijəhəs* 'species of sago' and *lajj* ~ *lajjəhəs* 'species of long yam'. They are evidently borrowings (the latter is an adaptation of Boikin *lanji*) and are assigned to class VIII in light of their plurals.

gories except *k*, regardless of whether the phoneme is typically a plural marker, and in some cases even irrespective of its plural. In other words, the simple recognition of a sound as such is in some cases sufficient to trigger phonological agreement on a form, in a sense disregarding the categorization of the lexeme as a whole (26c).

(26)	<u>Noun</u>	<u>Agreement</u>	
a.	<i>yabig^w</i>	<i>g^w</i>	‘soup’
	<i>kwapec</i>	<i>c</i>	‘many-pointed spear’
	<i>ñiburir</i>	<i>r</i>	‘vertical house post’
	<i>ucap</i>	<i>p</i>	‘undergrowth’
	<i>arəs</i>	<i>s</i>	‘sago bark bed’
b.	<i>ət-is arəs</i>		‘one bed’
	<i>bi-əs arəs</i>		‘two beds’
	<i>ət-ib oub</i>		‘one coconut’
	<i>worubai-bi oub</i>		‘many coconuts’
c.	<i>ət-u su</i>		‘one shoe’
	<i>bi-ou su</i>		‘two shoes’
	<i>ət-um brum</i>		‘one broom’
	<i>worubai-mi brum</i>		‘many brooms’
	<i>lam ma-tau urupat</i>		‘the lamp stands in the house’
	<i>bi-ec laməhəs (*lap)</i>		‘two lamps’

There are also cases in which agreement is idiosyncratically hyper-alliterative. An example from Cemaun is the noun *magus* ‘stove, earth base of fireplace’, which takes suffixal agreement with *-gus* (as opposed to the regular agreement suffix *-gəs*). Another is the noun *abər* ‘water’, which takes suffixal agreement with *-bər* (as opposed to the regular class II agreement suffix *-bir*). In Bukiyip, the borrowed noun (*moto*)*baik* ‘motorcycle’ is sometimes pronounced with *p*, instead of *k*: *motobaip*. In that case its final sound is projected in agreement forms with *p*: *na-pi-bikik* ‘VII.SG-p-put.away’ (Conrad n.d.:NB108).

A striking source of evidence for the phonological agreement rule comes from a phenomenon in Fortune’s Rohwim materials that involves agreement with an allomorph (indeed perhaps even an allophone) of a noun terminal. There are several classes that mark the plural with *s*. One and only one of these, class III, has an optional variant replacing plural *s* with a glottal stop. Some examples are given in (27).

- (27) *baijag* (SG) ~ *baijas* or *baija'* (PL) 'sago bark'
meitag (SG) ~ *meitagas* or *meitaga'* (PL) 'cassowary trap'
barawag (SG) ~ *barawas* or *barawa'* (PL) 'spear'

Glottal stop is at best quasi-phonemic in Rohwim; it appears in Fortune's transcriptions quite often, mostly as a realization of *k*, though Fortune insists that no distinctions in meaning ever hinge on its presence.¹⁵ Now consider how these glottal stop-final forms are treated in agreement. When the agreement marker is realized as a prefix or an infix, the regular agreement forms with *s* appear, as shown by the verbal prefix *gas* in (28a,b). However, when the agreement marker is word-final, it too is realized as glottal stop instead of *s*, as seen in (28c,d).

- (28) a. *ana-ga'* *pu-gas-a'* *mahiga'*
 INDEF-III.PL 2PL-III.PL-eat meat.III.PL
 'you all will eat some meat'
- b. *kwa-tuki* *barawa'* *kwa-gas-asahi*
 IV.SG-take.out spear.III.PL IV.SG-III.PL-shoulder
 'she takes out spears and carries them here on her shoulders'
- c. *barawa'* ... *akwokw-i-ga'*
 spear.III.PL IV.SG-POSS-III.PL
 'her... spears'
- d. *meitaga'* *bi-oga'*
 cassowary.trap.III.PL two-III.PL
 'two cassowary traps'

The glottal stop does not appear in place of *s* on an agreeing element unless it appears in place of *s* on a noun; nor are there any instances of *s* on

¹⁵Changes between *k* and the glottal stop can be made with impunity, as no two words which depend upon a *k* and a glottal stop distinction for their difference in meaning exist in the language. In practice, in [the dialect he describes], the *k* is in very general use, and the glottal stop substitute is only occasionally used instead of a terminal *k*, usually in a verb or in a pronoun' (Fortune 1942:4). An example of glottal stop substituting for *k* can be found in example (28a) above where the verb stem *-ak* 'eat' is pronounced *-a'*. Cemaun has nothing resembling this process; final *s* is always pronounced on class III nouns and their agreeing elements, and glottal stop is never substituted for *k* (though *k* frequently and in some contexts obligatorily deletes in word-final position). My 1998 visit to Woginara, the mountain settlement where traditional speakers of Fortune's dialect are now located, uncovered only a single instance of a process resembling glottal stop substitution for *s*. The controller is *ubas* 'feces', which is not obviously a class III noun, and the final *s* and its corresponding agreement markers are deleted, rather than debuccalized.

final agreeing elements when the noun itself is pluralized with glottal stop. In this case it is not possible for agreement to be parasitic on an already existing agreement form, since glottal stop does not mark agreement anywhere else in the Arapesh class system. It should also be emphasized that this phenomenon does not reflect an automatic phonetic process but is crucially sensitive to morphological information: only the class III plural *s*, and not the class I, class VIII, or class IX plural *s*, is subject to word-final glottal substitution. The confinement of glottal substitution to a single class further confirms that the various *s* plurals are not collapsed morphologically into a single category, so the development of the new agreement pattern for borrowed nouns with *s* is not simply assimilation to a dominant or unmarked plural form.

This entire array of observations is unified by the analysis presented in Table 1. For the sake of concreteness the specific rules are written for Rohwim, but an essentially parallel analysis could be stated for other Arapesh dialects. The table construes the system as a hierarchy of rules or constraints, and the resemblance to constraint hierarchies within approaches like optimality theory is not accidental. It should be borne in mind, however, that the hierarchy here is meant to establish the possible layers within the system rather than to directly derive surface forms. Thus, while these particular rules are language-specific, the hierarchical ordering of rules into blocks according to information type (semantic, morphological, and phonological) is intended as a general typological statement about noun class assignment systems.

SEMANTIC	MORPHOLOGICAL	PHONOLOGICAL
	REGULAR	DEFAULT
+MASCULINE → VII	{/Xbʲ/[+SG] ↔ /Xbʲys/[+PL]} → I	{/X/[+SG] ↔ /Xehas/[+PL]} → VIII
+FEMININE → IV	{/Xbøʀ/[+SG] ↔ /Xrʲyb/[+PL]} → II	/XC/ → AGRC
	{/Xg/[+SG] ↔ /Xs/[+PL]} → III	
	{/Xkʰ/[+SG] ↔ /XV _[+LAB] /[+PL]} → IV	
	{/Xm/[+SG] ↔ /Xpʲ/[+PL]} → V	
	{/Xn/[+SG] ↔ /Xb/[+PL]} → VI	
	{/Xn/[+SG] ↔ /Xm/[+PL]} → VII	
	{/Xñ/[+SG] ↔ /Xʃ/[+PL]} → VIII	
	{/Xpʰ/[+SG] ↔ /Xs/[+PL]} → IX	
	{/Xr/[+SG] ↔ /Xgʷ/[+PL]} → X	
	{/Xt/[+SG] ↔ /Xgʰ/[+PL]} → XI	
	{/Xuh/[+SG] ↔ /Xruh/[+PL]} → XII	
	{/XVh/[+SG] ↔ /XV _[+FR] h/[+PL]} → XIII	

Table 1. The Arapesh Noun Class Assignment System

SEMANTIC	MORPHOLOGICAL		PHONOLOGICAL
	REGULAR	DEFAULT	
<i>afuken</i> 'elder brother to a man'	+MASC → VII { <i>afuken</i> ~ <i>afukenim</i> } → VII		/afuken/ → AGR _n
<i>misis</i> 'European woman' (TP)	+FEM → IV		/misis/ → AGR _s
<i>nybat</i> 'dog'		{ <i>nybat</i> ~ <i>nybagu</i> } → XI	/nybat/ → AGR _t
<i>maliu</i> 'species of rattan'		{ <i>maliku</i> ~ <i>maliu</i> } → IV	/maliu/ → AGR _u
<i>pas</i> 'taro pounder'		{ <i>pas</i> ~ <i>pasehas</i> } → VIII	/pas/ → AGR _s
<i>kes</i> 'suitcase' (TP)			/kes/ → AGR _s

Table 2. The Classification of Selected Arapesh Nouns

The analysis says the following. Rules that assign class on the basis of nominal semantics have the first say. In Arapesh there are two such rules, and they have no exceptions. The class assignment of the borrowed human noun *misis* ‘European woman’ in Table 2 is determined entirely by semantic factors, allowing nouns to disregard the simultaneous pressures exerted by the form-based rules, a degree of freedom with consequences that were noted in Chapter 4. Even so, native human nouns are almost always convergently categorized by both semantics and morphology, as shown in the case of *afuken* ‘older brother to a man’. This redundancy surely adds to the system’s robustness.

For nouns lacking any class-relevant semantics such as *nybat* ‘dog’ in Table 2, the determination of class is made by the regular morphological assignment rules. These refer to information that is more abstract than an individual noun form; in Arapesh they refer to the singular-plural relation. Morphological assignment is at least partially redundant with the phonological assignment rules for normal singular nouns, and for most plurals (the table illustrates this with *maliu* ‘species of rattan’), though not for those exceptional nouns that take either the default plural *-ehas* or some other unpredictable plural. The key to understanding the assignment of the native Arapesh *s*-final noun *pas* ‘taro pounder’ in the table is to recognize that it also participates in the number marking system through the morphological default plural *-ehas*, and it is assigned to the morphological default category class VIII as a result.

Finally, when none of these rules has anything to say about a noun’s class, as in the case of new *s*-final nouns like *kes* ‘suitcase’ that do not have distinct plural forms at all, the decision falls to the simplest, most iconic determinant of class: reduplicative or alliterative assignment, which refers to a noun at the level of its concrete phonological form and on that basis extrapolates for it an entire paradigm of alliterative agreement markers. This last kind of default rule asserts itself only rarely—it is only rarely forced to—and it is regularly satisfied indirectly in any case by the working of the higher-level rules. But with the introduction of a large set of new nouns that are not readily interpretable by either of the two higher-priority rule types, the phonological default is set in motion. I suggest that what we are seeing in the treatment of *s*-final borrowings is a redundant generalization called forth from an uninfluential position in the back of the lexicon and brought into active, productive operation in the language. The result is agreement of a most natural, primal, or ‘universal’ kind (Aronoff 1998).

Of course, to propose that a phonological classification and agreement rule plays some role in Arapesh grammar is not at all the same thing as saying that it defines the entire Arapesh system or that it determines the agreement properties of every noun in the language—far from it. The existence

of non-alliterative agreement with lexical plurals of the kind illustrated in (15) above is accounted for in our analysis by the operation of ordinary morphological noun class assignment rules. To suggest, as Corbett 2006:90 does, that such examples ‘demonstrate that we do not have an alliterative agreement system’ is to misunderstand the claim put forth here. Arapesh makes use of an alliterative agreement rule, but it is not the language’s only agreement rule. Our analysis posits classification rules of multiple types, each of which occupies a well-defined place in the hierarchy of rules, and each of which exerts limited influence on the overall system. Systematic surveys of noun class systems find no language in which class does not have at least one semantic correlate (Corbett 1991, Aikhenvald 2000), and as we have seen, this generalization also holds also in Arapesh. What the analysis here adds is a distinction between the use of morphological and phonological information as a means of default. Phonological assignment can be seen as the ultimate last resort strategy, a rule type that is only called upon when standard morphological defaults are for some reason inapplicable. The analysis thus enriches the notion of defaults to include the possibility of opting out of mediation through organized morphological classes altogether.

Not only does the *s*-agreement phenomenon provide another source of evidence that phonological form can be directly relevant for class assignment, but it also reveals the degree to which the assignment of nouns to classes is bound up with the realization of class in agreement.¹⁶ Chapter 6 considers the consequences of this striking phenomenon of ‘literal alliterative concord’ for a theory of grammatical architecture.

5.3 A Contrastive Case: Noun Class Restructuring in Bantu

In Chapter 1, I tried to explain what was typologically unusual about Arapesh by drawing a contrast between the factors governing class assignment in the Abu’ dialect of Arapesh and those governing assignment in the Bantu language Swahili. Both have large noun class systems in which many classes show a significant correlation with nominal form. But whereas the Swahili noun classes are associated with a noun’s phonological form only indirectly by way of a limited set of noun prefixes, the Abu’ classes correspond much more directly to a phonological characteristic of noun form, the noun’s final phonological element. Importantly, the quality of noun-final elements is determined essentially freely according to Abu’ phonotactics.

¹⁶See Audring 2009 for an interesting effort to tie agreement to assignment at the other end of the typological spectrum, in small gender systems with pronominal agreement only.

By now it should be clear that a similar situation holds for the class-defining noun-final consonants in Rohwim Arapesh and other Arapesh varieties as well. It is this direct connection between phonological form and morphosyntactic categorization that seems to underlie the innovation of *s*-agreement with singulars in Bukiyip and Mufian.

Here we summarize the changes that have taken place in the noun class systems of three Bantu vernaculars, all of which have diverged in structure from the canonical models of their corresponding standard or lexifier varieties as a consequence of intensive language contact. The question is whether the situation has led any of these systems to extend the number of classes by innovating new prefixes, and whether they have developed alliterative agreement with new prefixes. To preview the answer, what we find is consistent with the fundamental difference we argue separates the Arapesh and Bantu systems in terms of their organization: in none of the Bantu varieties has noun classification been restructured in a way that suggests a direct connection between phonological form and morphosyntactic categorization.

The cases we will look at here are from three Bantu lingua francas that have been analyzed in detail by Bokamba 1993 and discussed from a socio-historical perspective by Mufwene 1989. The first of these is KiTuba, to be compared with the set of varieties known as ethnic KiKongo (Mufwene 1997). Following Bokamba 1993 and Mufwene 1997, ethnic KiKongo is represented by the composite 'diasystem' variety. The second is the urban dialect of LiNgala (as spoken in Kinshasa, Zaire and Brazzaville, Congo), which is treated in relation to standard literary LiNgala. The third variety we will consider is the dialect of Swahili spoken in Shaba Zaire, as described by Kapanga 1991, the standard counterpart of which is assumed to be Tanzanian coastal Swahili.¹⁷ The comparison of these languages with Arapesh is undertaken here in order to highlight formal changes in superficially similar systems that have responded to a situation of intensified contact, as Arapesh has with Tok Pisin, with an eye toward inferring the classification principles involved.

KiTuba and Kinshasa LiNgala share a common restructuring in the morphological realization of nouns. In KiKongo and standard LiNgala, nouns are characterized by conventional Bantu singular ~ plural prefix

¹⁷The history of each of these Bantu languages, and the relation of each to its corresponding standard or lexifier, varies. As mentioned above, standard KiKongo is a composite language, originally constructed to represent a group of closely related varieties. We also know that early standard LiNgala was embellished by missionaries (see Mufwene 1989:80-81). But what is clear in each case is that over time natural historical developments from an ethnic Bantu source have taken place.

pairs, as shown in (29) and (30). The data here is from Bokamba 1993 and preserves his transcription. Tones are not relevant to distinguishing classes and are not represented.

(29) KiKongo

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>mw-ana</i>	<i>ba-ana</i> [bana]	‘child’
1 ~ 2	<i>Ø-tata</i>	<i>ba-tata</i>	‘father’
3 ~ 4	<i>mu-sapi</i>	<i>mi-sapi</i>	‘finger’
3 ~ 4	<i>mw-inda</i>	<i>mi-inda</i> [minda]	‘light’
5 ~ 6	<i>di-kutu</i>	<i>ma-kutu</i>	‘ear’
7 ~ 8	<i>ki-baka</i>	<i>bi-baka</i>	‘wall’
7 ~ 8	<i>ki-ti</i>	<i>bi-ti</i>	‘chair’
9 ~ 10	<i>N-bwa</i> [mbwa]	<i>N-bwa</i> [mbwa]	‘dog’

(30) LiNgala

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>mw-ana</i>	<i>ba-na</i>	‘child’
1 ~ 2	<i>Ø-tata</i>	<i>ba-tata</i>	‘father’
3 ~ 4	<i>mu-nkanda</i>	<i>mi-nkanda</i>	‘book’, ‘letter’
5 ~ 6	<i>li-toko</i>	<i>ma-toko</i>	‘spoon’
7 ~ 8	<i>e-kuki</i>	<i>bi-kuki</i>	‘door’
9 ~ 10	<i>N-soso</i>	<i>N-soso</i>	‘chicken’

While these pairings are for the most part maintained in KiTuba and Kinshasa LiNgala, the class 2 plural prefix *ba-* has in both cases come to be added to nouns of other classes, as in (31) and (32), respectively.

(31) KiTuba

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>mw-ana</i>	<i>ba-ana</i> [bana]	‘child’
1 ~ 2	<i>Ø-tata</i>	<i>ba-tata</i>	‘father’
3 ~ 4	<i>mu-sapi</i>	<i>mi-sapi</i>	‘finger’
3 ~ (2-)4	<i>mw-inda</i>	<i>(ba-)mi-inda</i> [-minda]	‘light’
5 ~ 6	<i>di-kutu</i>	<i>ma-kutu</i>	‘ear’
7 ~ 8	<i>ki-baka</i>	<i>bi-baka</i>	‘wall’
7 ~ (2-)8	<i>ki-ti</i>	<i>ba-ki-ti</i>	‘chair’
9 ~ (2-)10	<i>N-bwa</i> [mbwa]	<i>ba-N-bwa</i> [mbwa]	‘dog’

(32) Kinshasa LiNgala

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>mw-ana</i>	<i>ba-na</i>	'child'
1 ~ 2	<i>Ø-tata</i>	<i>ba-tata</i>	'father'
3 ~ (2-)4	<i>mu-nkanda</i>	<i>(ba-)mi-nkanda</i>	'book', 'letter'
5 ~ (2-)6	<i>li-toko</i>	<i>(ba-)ma-toko</i>	'spoon'
7 ~ (2-)8	<i>e-kuki</i>	<i>(ba-)bi-kuki</i>	'door'
9 ~ 2-10	<i>N-soso</i>	<i>ba-N-soso</i>	'chicken'

In KiTuba the generalization of *ba-* is limited to classes 8, 10, and optionally 4, but in Kinshasa LiNgala the class 2 prefix is optionally present on nouns of every class 'with the [double-prefixed plurals] being more common' (Bokamba 1993:215). In its traditional class 2 environment the *ba-* plural is reserved for human nouns, but with its spread to other classes this semantic condition is being lost. Note also that in the innovative cases plural *ba-* is attached neither to the singular prefix-stem combination, nor to the noun stem alone, but rather to an already pluralized noun, creating a situation in which plurality is double-marked (albeit optionally) in every plural class in the system.

Differences in the realization of Shaba Swahili nouns relative to standard Swahili are much less dramatic. The basic noun paradigms are presented in (33) and (34).

(33) Standard Swahili

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>m-toto</i>	<i>wa-toto</i>	'child'
3 ~ 4	<i>m-sumari</i>	<i>mi-sumari</i>	'nail'
5 ~ 6	<i>ji-cho</i>	<i>ma-cho</i>	'eye'
7 ~ 8	<i>ki-su</i>	<i>vi-su</i>	'knife'
9 ~ 10	<i>N-umba</i> [nyumba]	<i>N-umba</i> [nyumba]	'house'
11 ~ 10	<i>u-shanga</i>	<i>N-shanga</i> [shanga]	'bead'

(34) Shaba Swahili

	<u>Singular</u>	<u>Plural</u>	
1 ~ 2	<i>mu-toto</i>	<i>ba-toto</i>	'child'
3 ~ 4	<i>mu-sumari</i>	<i>mi-sumari</i>	'nail'
5 ~ 6	<i>ri-cho</i>	<i>ma-cho</i>	'eye'
7 ~ 8	<i>ki-su</i>	<i>bi-su</i>	'knife'
9 ~ 6-10	<i>N-umba</i> [nyumba]	<i>ma-N-umba</i>	'house'
11 ~ 11	<i>bu-shanga</i>	<i>bu-shanga</i>	'bead'

In Shaba Swahili we see a number of shifts in the form of the prefixes, such as the class 8 plural *bi-* in place of standard Swahili *vi-*, *ba-* in place of *wa-*, and *bu-* in place of *u-*, but these do not alter the overall system. The changes are clearly sensitive to the word's morphological structure; as Kapanga 1991:170 notes, for example, the initial sound in *vimba* 'swell' is preserved (**bimba*) because it is not a noun class prefix in this word. There has, however, been a morphologically significant change affecting nouns of class 9/10, which in Standard Swahili are identical in the singular and plural: in the Shaba dialect the plurals are additionally prefixed with the class 6 marker *ma-*. The reanalysis of this class, which serves to provide it with overt marking for number, is common in Central Bantu languages (Bokamba 1993:216).

These brief sketches reveal two Bantu prefixes, class 2 *ba-* and class 6 *ma-*, which appear to be expanding their distribution in the way characteristic of morphological defaults. Were the default plural of Arapesh behaving similarly, we would find the class VIII plural *-ehas* suffixed to nouns of other classes, producing alternations along the lines of $/Xn/_{[+SG]} \sim /Xb(-ehas)/_{[+PL]}$ or $/Xp/_{[+SG]} \sim /Xs(-ehas)/_{[+PL]}$. But there is no evidence that *-ehas* is spreading in this way. There are some exceptional nouns that take *-ehas* rather than their expected plural, and *-ehas* is occasionally found as a secondary plural marker creating alternations like Cemaun *keñ* \sim *kecəhəs* 'bow', *cewebam* \sim *cewebapəhəs* 'species of lizard', and the Tok Pisin borrowing *kap* \sim *kasəhəs* 'cup'. But neither *-əhəs* nor any other plural has been extended to the point of subsuming an entire class, as has occurred in each of these Bantu varieties.

Do we have any evidence that agreement has shifted in ways that correspond to the shifts in noun realization in these languages? Bokamba 1993:217-218 frames the question nicely:

Since it is traditionally assumed in Bantu linguistics that the classification of noun prefixes into singular/plural pairs is motivated by grammatical agreement properties, it would be informative to examine... whether there is a direct correlation between the singular/plural pairings established by the noun morphology and the grammatical agreement patterns they trigger on verbs and modifiers in the syntax. [D]o morphological noun classes... entail syntactic noun classes in these languages?

Assuming that the Bantu classes are mediated by morphological assignment rules rather than being determined by phonological form directly, new alliterative agreement classes would not be expected to arise.

KiKongo exhibits the substantially alliterative agreement pattern typical of other Bantu languages, as shown in (35).

(35) KiKongo

<i>mwa-ana</i>	<i>wu-w-a-mbote</i>	<i>wu-bwi-idi</i>
1-child	this-1-CNCT-good	1-fall-PRF
‘this good child has fallen’		
<i>ba-ana</i>	<i>ba-ba-mbote</i>	<i>ba-bwi-idi</i>
2-child	this-2-CNCT-good	2-fall-PRF
‘these good children have fallen’		
<i>ki-ti</i>	<i>ki-ki-a-mbote</i>	<i>ki-tolok-ele</i>
7-chair	this-7-CNCT-good	7-break-PRF
‘this good chair has broken’		
<i>bi-(ki-)ti</i>	<i>bi-bi-a-mbote</i>	<i>bi-tolok-ele</i>
8-(7-)chair	this-8-CNCT-good	8-break-PRF
‘this good chair has broken’		

The KiTuba system is significantly restructured relative to KiKongo, as shown in (36). KiTuba has no class agreement at all, not even for animates. The nouns themselves continue to be overtly marked for class, but those classes are no longer of any syntactic significance.

(36) KiTuba

<i>mw-ana</i>	<i>ya</i>	<i>mbote</i>	<i>yayi</i>	<i>mene</i>	<i>bwa</i>
1-child	CNCT	good	this	PRF	fall
‘this good child has fallen’					
<i>b-ana</i>	<i>ya</i>	<i>mbote</i>	<i>yayi</i>	<i>mene</i>	<i>bwa</i>
2-child	CNCT	good	this	PRF	fall
‘these good children have fallen’					
<i>ki-ti</i>	<i>ya</i>	<i>mbote</i>	<i>yayi</i>	<i>mene</i>	<i>bukana</i>
7-chair	CNCT	good	this	PRF	broken
‘this good chair has broken’					
<i>ba-ki-ti</i>	<i>ya</i>	<i>mbote</i>	<i>yayi</i>	<i>mene</i>	<i>bukana</i>
2-8-chair	CNCT	good	this	PRF	broken
‘these good chairs have broken’					

The restructuring that has taken place in Kinshasa LiNgala is also quite interesting. In contrast to the full verbal agreement paradigm of the standard variety, examples from which are shown in (37), noun class agreement in Kinshasa LiNgala is reduced to two categories: a human class differentiated for singular and plural, and everything else irrespective of number (38).

(37) Standard LiNgala

mw-ana *a-kobunga*
 1-child 1-will.get.lost
 ‘the child will get lost’

ba-ana *ba-kobunga*
 2-child 2-will.get.lost
 ‘the children will get lost’

mu-ndele *mu-kobunga*
 3-white person 3-will.get.lost
 ‘the white person will get lost’

mi-ndele *mi-kobunga*
 4-white person 4-will.get.lost
 ‘the white people will get lost’

mu-nkanda *mu-kobunga*
 3-book/letter 3-will.get.lost
 ‘the book/letter will get lost’

mi-nkanda *mi-kobunga*
 4-book/letter 4-will.get.lost
 ‘the books/letters will get lost’

e-kuki *e-kobunga*
 7-door 7/9-will.get.lost
 ‘the door will get lost’

bi-kuki *bi-kobunga*
 8-door 8-will.get.lost
 ‘the doors will get lost’

(38) Kinshasa LiNgala

mw-ana *a-kobunga*
 1-child 1-will.get.lost
 ‘the child will get lost’

ba-ana *ba-kobunga*
 2-child 2-will.get.lost
 ‘the children will get lost’

mu-ndele *a-kobunga*
 3-white person 3-will.get.lost
 ‘the white person will get lost’

<i>ba-mi-nde</i>	<i>ba-kobunga</i>
2-4-white person	4-will.get.lost
'the white people will get lost'	
<i>mu-nkanda</i>	<i>e-kobunga</i>
3-book/letter	7/9-will.get.lost
'the book/letter will get lost'	
<i>ba-mi-nkanda</i>	<i>e-kobunga</i>
2-4-book/letter	7/9-will.get.lost
'the books/letters will get lost'	
<i>e-kuki</i>	<i>e-kobunga</i>
7-door	7/9-will.get.lost
'the door will get lost'	
<i>ba-bi-kuki</i>	<i>e-kobunga</i>
2-8-door	7/9-will.get.lost
'the doors will get lost'	

Whereas the form that has spread as a secondary plural on Kinshasa LiNgala nouns is the class 2 prefix *ba-*, the form that has replaced the agreement markers for all non-human classes is class 7/9 *e-*. The historical markers still appear on nouns; i.e., the prefixes on human nouns remain essentially unchanged. But Kinshasa syntax is unable to distinguish among nouns apart from their semantics. The language thus evidences an uncoupling of class from morphological realization and from form more generally. Agreement has become much less alliterative as a result, with a constant verbal subject agreement marker now extended to all non-human nouns irrespective of their form (and adjective agreement is lost altogether). The changes that have occurred in KiTuba and Kinshasa LiNgala have effectively reduced the noun classes to a series of lexical stem sets along the lines of English *leaf* ~ *leaves*, but here pervading the entire nominal vocabulary. In this way Kinshasa LiNgala is just like Weri: a full array of realization classes is maintained for nonhuman nouns, but they have no function beyond the morphological marking of number.

Developments in the Shaba Swahili agreement system are more complex. In standard Swahili agreement with modifying adjectives is for the most part alliterative, whereas demonstrative and verbal subject agreement is much less so (39):

(39) Standard Swahili

<i>mu-toto</i>	<i>yu-le</i>	<i>m-dogo</i>	<i>a-me-potea</i>
1-child	1-that	1-small	1-PRF-lose
‘that small child is lost’			
<i>wa-toto</i>	<i>wa-le</i>	<i>wa-dogo</i>	<i>wa-me-potea</i>
2-child	2-that	2-small	2-PRF-lose
‘those small children are lost’			
<i>m-sumari</i>	<i>u-le</i>	<i>m-dogo</i>	<i>u-me-potea</i>
3-nail	3-that	3-small	3-PRF-lose
‘that small nail is lost’			
<i>Ø-jembe</i>	<i>li-le</i>	<i>Ø-dogo</i>	<i>li-me-potea</i>
5-hoe	5-that	5-small	5-PRF-lose
‘that small hoe is lost’			
<i>ma-jembe</i>	<i>ya-le</i>	<i>ma-dogo</i>	<i>ya-me-potea</i>
6-hoe	6-that	6-small	6-PRF-lose
‘those small hoes are lost’			
<i>ndoo</i>	<i>i-le</i>	<i>n-dogo</i>	<i>i-me-potea</i>
9-pail	9-that	9-small	9-PRF-lose
‘that small pail is lost’			
<i>u-fagio</i>	<i>u-le</i>	<i>m-dogo</i>	<i>u-me-potea</i>
11-broom	11-that	1/3-small	11-PRF-lose
‘that small broom is lost’			

Bokamba 1993:231 describes the Shaba variety as simplified with respect to standard Swahili and as having an ‘essentially alliterative agreement system.’ However, a careful look at the agreement data reveals a less straightforward profile. Examples are given in (40).

(40) Shaba Swahili

<i>mu-toto</i>	<i>u-le</i>	<i>mu-zuri</i>	<i>a-ri-anguka</i>
1-child	1-that	1-nice	1-PST-fall
‘that nice child fell’			
<i>ba-toto</i>	<i>ba-le</i>	<i>ba-zuri</i>	<i>ba-ri-anguka</i>
2-child	2-that	2-nice	2-PST-fall
‘those nice children fell’			

<i>mu-fuko</i>	<i>i-le</i>	<i>mu-zuri</i>	<i>i-ri-anguka</i>
3-bag	3-that	3-nice	3-PST-fall
‘that nice bag fell’			
<i>ri-papayi</i>	<i>ri-le</i>	<i>ri-zuri</i>	<i>ri-ri-anguka</i>
5-papaya	5-that	5-nice	5-PST-fall
‘that nice papaya fell’			
<i>ma-papayi</i>	<i>i-le</i>	<i>mu-zuri</i>	<i>i-ri-anguka</i>
6-papaya	6-that	6-nice	6-PST-fall
‘those nice papayas fell’			
<i>nyumba</i>	<i>i-li</i>	<i>mu-zuri</i>	<i>i-ri-anguka</i>
9-house	6-that	6-nice	6-PST-fall
‘that nice house fell’			
<i>lu-nsonga</i>	<i>lu-le</i>	<i>lu-zuri</i>	<i>lu-ri-anguka</i>
20-plant	20-that	20-nice	20-PST-fall
‘that nice plant fell over’			

Beyond class 1, which derives its coherence primarily from the semantic correlate +HUMAN, Shaba Swahili has spread a constant form of the adjective prefix *mu-* to several other classes. The spread of this form is checked by the simultaneous extension of alliterative forms across the agreement series, as illustrated for example by classes 2 and 5.¹⁸ The appearance of class 20 is another divergence from standard Swahili.¹⁹ This class includes nouns prefixed with *lu-*, and as shown above, takes alliterative agreement with the prefix.

Changes are also found in the nouns used in locative phrases. In Standard Swahili, nouns are marked with a constant locative suffix *-ni*. There are three semantically differentiated agreement forms: *pa-* for a specific location, *mu-* for a specific location inside something, and unmarked *ku-*. In Shaba Swahili, the nouns themselves are marked with these prefixes and not with *-ni*, bringing the noun into formal alignment with the agreeing elements; compare (41) with (42).

¹⁸I have not presented examples from noun classes that are alliterative in both Swahili varieties since our interest is in agreement developments that increase alliteration.

¹⁹I follow Kapanga 1991 in using the term ‘class 20’ as a label for these nouns.

(41) Standard Swahili

nyumba-ni pa-le pa-zuri pa-mejaa na watu
 house-LOC LOC-that LOC-nice LOC-is.full with people
 ‘that nice place at the house is full of people’

(42) Shaba Swahili

*pa-nyumba pa-le pa-zuri pa-napendeza*²⁰
 LOC-house LOC-that LOC-nice LOC-is.pleasing
 ‘that nice place at the house is pleasing’

Of all the changes considered in these Bantu contact languages, the extension of locative agreement markers to nouns and the development of a new *lu*-prefixed class represent a move toward increased alliteration, and might therefore be taken as evidence for the activation of a phonological principle in Swahili. As it turns out, however, in neither of these situations is the development motivated by pressure from the standard Swahili system. Rather, a model for each of the alliterative patterns can be found in other local language varieties.

According to Kapanga 1991:174, ‘the nominal morphological changes perceived in [Shaba Swahili] constitute an adoption of the morphological rendering of the noun class system of the local Bantu languages of Shaba.’ Locative nouns are prefixed with *ku-*, *pa-*, and *mu-* not only in Shaba Swahili, but also in the local Shabian languages Luba-Kasai, Bemba, Lunda, and Luba-Shaba (Kapanga 1991:173-174), from which the Shaba Swahili patterns apparently were adopted. A noun class encompassing nouns prefixed with *lu-* (and taking alliterative agreement with *lu-*; see Spitulnik 1987:128 for the noun class paradigm from ChiBemba) similarly exists in the local contact languages Luba-Kasai, Bemba, and Luba-Shaba, and it is this class that appears to be exerting an influence on Shaba Swahili. So, for example, the Shaba Swahili noun *lu-rimi* ‘tongue’ (see Kapanga 1991:171) has a *lu*-prefixed cognate in ChiBemba (*u-lu-limi*; see Spitulnik 1987:72), although its standard Swahili cognate is classified in class 11/10 (*u-limi*; see Ashton 1944:106).

In sum, then, the contact-induced changes that have taken place in the three Bantu varieties discussed here involve restructuring in the morphological realization of nouns (e.g., the spread of *ba-* to all plurals in Kinshasa LiNgala) and the loss of morphosyntactic distinctions among nouns (e.g., the concomitant neutralization of agreement marking to a single class for all

²⁰This example is extrapolated from the charts and discussion in Kapanga 1991 and Bokamba 1993.

inanimates irrespective of number). In no case, however, has classification been extended in accordance with a phonological principle. In the one instance in which it might appear that a new class has been innovated along with a new set of agreeing elements (the Shaba Swahili *lu*-class), the class turns out to be borrowed wholesale from another language.

The evidence for phonological classification from the assignment of borrowed nouns in Bantu is mixed. In a study dedicated to understanding the effect of borrowings on the Swahili noun class system, Zawawi 1979:130 finds that

words of non-Bantu origin are not automatically assigned to a ‘class’ whose prefix indicator resembles their initial sound. Not all non-Bantu words with the same initial sound are found in the same ‘class,’ and not every non-Bantu word has only one pattern of concordial agreement. The indicators acquired by these words are semantically determined.

Emphasizing the central role of a morphological default class in Swahili loan classification, Kapanga 1991:266 notes that

there is a general tendency in Swahili to use [the nasal-prefixed class 9] as a catch-all class, in that any new word in the language which does not flaunt an overt noun class similar to the noun classes of Swahili will be put in class 9. Most borrowings in the language are grouped in this class.

In ChiBemba, both phonological and semantic factors influence loanword classification, though it appears ‘that 9/6 [*iN-/ama-*] and 5/6 [*i-/ama-*] have become the dominant class assignments for “non-personal” count nouns’, in part because they minimally distort the form of the loanword, and in part because they are minimally marked semantically (Spitulnik 1987:97). Again we find that Bantu noun class markers are typically prefixed onto newly introduced nouns as they are adapted into the system, whereas such adjustments are never made in order to accommodate borrowings into Arapesh (they are only assigned a plural). Thus, in cases where Bantu class assignment rules do seem to refer to a borrowed noun’s form, the rules are referring to morphological information (the prefix) that is created when the noun is reinterpreted as morphologically complex.

As a final point, we can note that the independence of morphological class marking on nouns from the syntactic ramification of class in agreement is a characteristic not only of Bantu but of the Niger-Congo family in general. A comparative survey of Niger-Congo languages along the spectrum from those with full-blown noun class systems characterized by extensive noun marking and agreement, through those of the ‘Kwa’-type that have lost their noun classes all or in part (Good 2012), shows numerous examples in which concord is preserved while class marking is lost, and conversely in which class markers are preserved when they are no longer bolstered by agreement. Only in the Kru languages (discussed in the next

chapter) do we see phonological agreement emerge: instead of being lost as a result of phonological attrition, the class suffixes are compressed into the noun stem where they predict both class and agreement forms. But while Bantu noun-class assignment and alliterative agreement may appear superficially similar to Arapesh, these form-based systems differ qualitatively. In Bantu, the noun class markers are morphological units that happen to be realized by certain sounds. In Arapesh, the morphological units and the sounds that realize them can be referenced by the grammar as one and the same thing.

Theoretical Consequences of Alliterative Concord

6.0 Introduction

It has long been a priority in linguistic theory to develop strong, substantive principles governing the way grammatical components can interact. One of the most general and influential such principles remains the Lexicalist Hypothesis, which prohibits syntactic access to word-internal structure. Very much in the lexicalist spirit, Pullum and Zwicky 1988, Zwicky 1984, 1987, 1992, and Zwicky and Pullum 1983, 1986a,b defend a principle of ‘phonology-free syntax’ (PPFS), stating that no syntactic rule can be subject to language-particular phonological conditions or constraints:

[S]yntax can be sensitive to abstract properties realized in the distribution of phonological features, but not to the specific phonological features. Though the conditions in a syntactic rule can have certain sorts of indirect or ultimate phonological consequences..., these conditions never seem to distribute phonological properties directly; no language has a syntactic rule stipulating that some constituent begin with an obstruent or have no more than two syllables, or contain only unrounded vowels, or have stress on its penultimate syllable’ (Zwicky 1996:301).

Many potential counterexamples to the PPFS have been proposed in the literature and more or less successfully deflected. Miller, Pullum, and Zwicky 1997:68 call the PPFS ‘the best and clearest example of a universal [law] of linguistic structure that has been uncovered by modern linguistics.’ Similar statements are made in Corbett 2006:89 and 2010:18.

In this chapter, we consider some of the empirical arguments against the PFS and present data from Kru languages, the Arapesh language Abu', and the West Atlantic language Bainuk in which an ostensibly morphosyntactic agreement operation makes reference to the phonological form of a word. We already saw evidence of such a 'literal' agreement phenomenon in the replacement of plural *-s* with glottal stop on class III nouns and agreement suffixes in Rohwim; we also saw it in the development of *s*-agreement in Mufian and Bukiyip, which was discussed in Chapter 5 from the point of view of class assignment. The further literal alliterative concordial patterns discussed here provide strong evidence that grammatical components can interact in a significantly freer way than is traditionally assumed in linguistic theory.

6.1 The Relationship Between Word Form and Syntax

At least since Chomsky's *Remarks on Nominalizations* (1970), which paved the way for distinguishing morphology from syntax as separate generative components in grammar, linguists have confronted the problem of how to define the lower boundary of syntax. The consequence has been a series of proposals all falling into the general category of 'lexicalism' or the 'lexicalist hypothesis'. The earliest versions of the lexicalist hypothesis essentially prohibited syntactic transformations from performing derivational morphology, so that, for example, the English nominalization *destruction* could not be related to its base verb *destroy* by a rule of syntax. Many revisions of the lexicalist hypothesis have since been formulated, e.g., the 'extended lexicalist hypothesis' (Jackendoff 1972), the 'generalized lexicalist hypothesis' (Lapointe 1988), and 'syntactic atomativity' (DiSciullo and Williams 1987). Despite their differences, all lexicalist proposals represent an effort to make precise the basic idea that '[t]he syntax neither manipulates nor has access to the internal form of words' (Anderson 1992:84).

Perhaps the most pressing problem for any lexicalist theory is how to deal with inflectional morphology, an area in which syntax and word-form inevitably interact. Anderson 1988, 1992 in fact defines inflection as just the ways in which the lexicalist hypothesis appears to be violated; that is, as just those facts about words that are syntactically relevant or syntactically visible. The key to making such a solution more than terminological consists in carefully defining the medium through which word-internal information is made available to syntax, so that violations of word structure occur only in limited ways. Otherwise, syntax could refer to virtually anything inside a word, leaving the definition of inflection theoretically empty and the lexicalist hypothesis unfalsifiable.

Anderson (1992:90) proposes that the medium of interface is a word's 'morphosyntactic representation' (MSR), the set of features that expresses a word's morphological properties, and that 'the only way syntax can affect the form of a word is by manipulating its MSR.' Such a requirement reformulates within a processual framework the 'atomicity thesis' couched in standard morphemic terms: '[an] affix can register a syntactic effect only by affecting the categorial status... of a word in a sentence' (DiSciullo and Williams 1987:48). The essential idea of the MSR is that what is syntactically significant about words are the bundles of morphological features that are associated with them. These features are drawn from a finite, specifiable set, whether universal or language-specific. Importantly, a word's morphosyntactic representation 'can be given a formal character that abstracts away from details of just what phonological material appears in the word and where it appears relative to other material' (Anderson 1992:89-90). Conversely, the phonological features that correspond to morphosyntactic features should not be syntactically visible. If they were, Anderson argues, we would expect these aspects of word form to be systematically available to syntactic manipulation. To the extent that phonological distinctions per se have no consequences beyond their role in realizing the word itself, there is justification for prohibiting the syntax from having access to phonological aspects of word-internal form.

A series of articles by Pullum, Zwicky, and others have advanced this suggestion to a theoretical principle, the 'principle of phonology-free syntax' or PPFS which states that syntactic rules are not subject to language-particular phonological conditions or constraints. The PPFS is probably the weakest form that a lexicalist statement could take: even if syntax can see word-internal morphological structure, at the very least it cannot see its phonological realization.

In this chapter I question the lexicalist claim that syntax is always phonology-free by considering cases of grammatical agreement, in which the form of one word depends on features of another in some grammatical context. According to the PPFS, no syntactic agreement rule should be able to make the form of one word depend on phonological features of another. We have already seen Arapesh data in which precisely this situation seems to obtain, and other compelling cases will be introduced here. The result is an insidious leak in the architectural position laid out above, according to which syntax manipulates only morphosyntactic representations, and phonological form is systematically invisible to inflectional processes. We will conclude that either phonological form must indeed be syntactically available in some circumstances, thus violating even the weakest possible form of the lexicalist hypothesis, or else morphosyntactic representations may be

transparent or dependent upon phonological form, thus weakening their crucial role as the medium of interface between word form and syntax.

6.2. Two Proposed PPFs Violations in Agreement

6.2.1 Somali

The first case we will consider is irregular verb agreement in the Cushitic language Somali (Hetzron 1972, Zwicky and Pullum 1983). Somali has two classes, masculine and feminine, which are distinguished only in the singular, and verbs agree with their subjects for number and class. The regular noun plural is *-(C)o*, which is subject to assimilation. As shown in (1) below, the agreement form of the verb ‘left’ is *tegay* with the masculine singular noun ‘truck’ (1a), *tegtay* with the feminine singular noun ‘woman’ (1c), and *tegeen* with plurals of either class (1b,1d). Hetzron 1972 suggests that Somali syntax makes reference to phonological form because a verb agreeing with a plural noun that is formally irregular is optionally realized in the form appropriate for agreement with a feminine singular subject, in addition to taking the normal plural agreement form. The plural of the irregular noun *dibí* ‘ox’ (1f) takes either the verbal agreement form *tegeen*, the normal plural agreement form, or more preferably *tegtay*, which is identical to the form used with feminine singulars.¹

- | | | | |
|--------|-----------------------------------|--------------------|--|
| (1) a. | <i>baabùur-kii</i>
truck-the | <i>wùu</i>
he | <i>tegay</i>
left (M.SG) |
| b. | <i>baabuurá-dii</i>
trucks-the | <i>wày</i>
they | <i>tegeen</i>
left (PL) |
| c. | <i>náag-tii</i>
woman-the | <i>wày</i>
she | <i>tegtay</i>
left (F.SG) |
| d. | <i>naagí-hii</i>
women-the | <i>wày</i>
they | <i>tegeen</i>
left (PL) |
| e. | <i>díbi-gii</i>
ox-the | <i>wùu</i>
he | <i>tegay</i>
left (M.SG) |
| f. | <i>díbí-dii</i>
oxen-the | <i>wày</i>
they | <i>tegeen ~ tegtay</i>
left (PL) ~ (F.SG) |

¹Nouns taking feminine agreement in the plural are often animate, though ‘not all of them and not only them’ (Hetzron 1972:261 n.5). For example, the plurals *ughán* ‘eggs’ and *banaadiix* ‘rifles’ also take feminine singular agreement.

One clue to what underlies the Somali agreement pattern with irregulars is the fact that their second syllable tonal accent is typical of feminine nouns, as shown in (2). It would thus appear that Somali speakers are producing verbs that agreement forms based on the way the nouns sound, rather than on the basis of their abstract morphosyntactic class features.

- (2) *ʔinan* ‘son’
 ʔinán ‘daughter’

However, reference to phonological form per se is insufficient to explain the odd Somali agreement pattern, since irregular plurals of other sorts take optional feminine singular agreement as well. For example, words with borrowed Arabic endings, as in (3a), and words with Arabic broken plural forms, as in (3b), also take the optional agreement pattern, even though such irregular nouns do not necessarily ‘sound feminine’, or apparently even foreign: ‘[Nouns like *nijàar* ~ *nijaarín*] cannot be regarded as unasimilated borrowings; they are phonologically impeccable in Somali’ (Zwicky and Pullum 1983:395).

- (3) a. *nijàar-kii* *wùu* *tegay*
 carpenter-the he left (M.SG)
 nijaarín-tii *wày* *tegeen* ~ *tegtay*
 carpenters-the they left (PL) ~ (F.SG)
- b. *màrkab* ‘ship’
 maraakíib ‘ships’

In accounting for these facts, Zwicky and Pullum 1983 argue that the Somali optional agreement rule refers not to a noun’s phonological form, but instead to the simple fact of its morphological irregularity, which it does via the abstract intrinsic feature [+Subplural]: ‘Verbs in finite clauses having [+Subplural, +Plural] subject NPs are optionally marked [+Feminine -Plural]’ (Zwicky and Pullum 1983:392). As a result, it is unnecessary for syntactic agreement in Somali to refer to the phonological form of the head noun directly, and morphosyntactic representations can be maintained as the sole medium through which morphology and syntax interact.

6.2.2 Vata

Reference to phonological form in the pronominal agreement system of the Kru language Vata is much less readily dismissed. Vata distinguishes five series of pronouns, one corresponding to each of the language’s five basic

vowel qualities. Which series an agreeing pronoun is drawn from depends upon the quality of the noun-final vowel, abstracting away from its tone and its value for the vocalic feature ATR (Kaye 1981, Zwicky 1987). The pattern is represented in (4) by the form of the 3rd person relative pronoun for each series.

(4)	<u>Series</u>	<u>Noun</u>	<u>Gloss</u>	<u>'be big'</u>	<u>Relative Pronoun</u>
	A	<i>ɟla</i> <i>sɬa</i>	lion house	<i>a-yli</i>	<i>mama</i>
	E	<i>cice</i> <i>ble</i>	eagle cow	<i>ɛ-yli</i>	<i>meme</i>
	I	<i>li</i> <i>bɔgi</i>	songs feet	<i>ɪ-yli</i>	<i>mimi</i>
	O	<i>lagɔ</i> <i>deto</i>	god spider	<i>ɔ-yli</i>	<i>mɔmɔ</i>
	U	<i>gɔlɔ</i> <i>su</i>	pirogue tree	<i>ɔ-yli</i>	<i>mɔmɔ</i>

As seen in (4), nouns ending in a low vowel take pronouns from the A-series, nouns ending in a mid-front vowel take pronouns from the E-series, and so on throughout the word-final vocalism. The system differs from Arapesh in distinguishing vowels rather than consonants, but this agreement pattern is still a classic instance of alliterative concord (Corbett 1991, Carstairs-McCarthy 1994) since 'the noun itself includes a form which is identical to the class agreement marker and which clearly indicates the class of the noun' (Corbett 1991:117). Alliterative concord involves an overt expression of class or noun class on a noun itself in conjunction with agreement forms that match the noun's form phonologically. A similar pattern is exhibited in agreement with third-person, non-human singular nouns in the Gbobo dialect of Krahn, a Kru language of Liberia and Côte d'Ivoire, where a five-way contrast among noun-final vowel qualities (ignoring tone and ATR) is reduced to three phonologically defined classes: front, high back, and non-high back, each of which takes agreement with matching vowels. If the Gbobo classes 'were the remnants of a noun class system... one would expect some evidence of a preponderance of certain types of nouns, such as those depicting liquids, large or small animals, or important or unimportant things' in some of the classes (Bing 1987:55-56), but as in Arapesh, this does not occur. The purely phonological nature of the Gbobo alliterative agreement rule is confirmed by its sensitivity to phonological

variation. An example is the pair *gbāé* ~ *gbāá* ‘stick’. The former variant is used by older speakers and receives front agreement following its final form; the latter is used by younger speakers with non-high back agreement, again following its final form.

Zwicky’s rules for Vata are presented in (5) below. Rule (5a) expresses the generalization that words ending in a vowel of a given height and backness belong to a certain class; rule (5b) illustrates this generalization in the specific case of a high back noun-final vowel. Rule (5c) expresses the distinct and overriding generalization that semantics takes precedence in situations of conflict: nouns denoting humans fall into the O-class regardless of which vowel they end with.

- (5) a. $XV_{[\alpha \text{ BK}, \beta \text{ HI}]} \# \rightarrow \text{class X}$
 b. /Xu/ \rightarrow U-class
 c. +HUMAN \rightarrow O-class (*ñema* ‘friend’, *kofi* ‘Kofi’, *jegle* ‘waiter’)

If agreement rules (5a) and (5b) are rules of inflectional morphology and thus involve syntactic operations, they are in violation of the PPFS, because they make direct reference to phonological form. But Zwicky concludes that the Vata pronominal agreement rules are not syntactic. Instead, he argues that these rules are implicational principles or redundancy rules which predict morphological features from phonological ones in the lexicon. In Zwicky’s view, the crucial point for establishing the lexical nature of these rules is that the classes they define are general nominal classes, rather than pronoun classes directly. Thus, what is transferred or matched in pronominal agreement in Vata is an abstract morphological class feature such as [+U-class], and not the noun-final high back vowel itself Zwicky 1987:5:

[T]he generalizations relating the phonological shapes of nouns and pronouns *can* be captured in an analysis using arbitrary features, via lexical redundancy rules... A familiar, quite general, principle of feature agreement then requires, at least in the default case, that anaphors have the same morphological features as their antecedents.

In other words, according to the lexical analysis, the Vata classes are not defined by morphosyntactic features, but rather by what Zwicky assumes are purely morphological features. This distinction is central to the analysis because, it will be recalled, the morphosyntactic representation is precisely the aspect of a word to which syntactic rules should have access. To allow morphosyntactic features to be determined directly by phonological form would make the lexical analysis empirically indistinguishable from a syntactic analysis; it would merely embed an additional step.

For Zwicky, the reason the Vata agreement rules are not strictly speaking syntactic is that pronominal agreement is ‘anaphoric’ as opposed to ‘local’. For purposes of evaluating the PPFS, the primary differences between anaphor and local agreement concern where the trigger and target are located with respect to one another, what kinds of targets are involved, and what kinds of features are involved. The differences are summarized in (6).

(6) <u>Anaphor Agreement</u>	<u>Local Agreement</u>
Trigger may be outside S	Trigger and target must be in same S
Morphosyntactic Features	Morphosyntactic Features
Purely Morphological Features	—
Proforms only	Proforms, Modifiers, Verb-Subject
Coreferential	Grammatical
Subject to Pragmatic Override	Resistant to Pragmatic Override

In anaphor agreement, the noun phrase trigger and the pronominal target co-occur within a discourse context, whereas in local agreement, trigger and target must occur within a syntactically specifiable domain; that is, they must appear within the same sentence. Anaphor agreement targets are proforms, whereas local agreement targets include proforms, modifiers, and verbs. As for the features they refer to, local and anaphor agreement may both refer to morphosyntactic features; for nouns that includes number and class or just class, as well as configurationally acquired properties such as case. But idiosyncratic, purely morphological features can only participate in anaphor agreement. How to determine which type of feature is involved in any given instance of agreement is not always obvious; however, if agreement is of the local type, then the relevant features are necessarily morphosyntactic. Since as far as we know Vata agreement involves only pronouns, Zwicky assumes that the relevant agreement features are purely morphological. Therefore, the fact that they are determined by phonological aspects of a noun causes no violation of the PPFS.

But the fact that local agreement refers only to morphosyntactic features does not necessarily tell us what features are involved in any given instance of anaphor agreement. There is also reason to assume that the control of anaphoric pronouns by their antecedents simply constitutes a subtype of a unified agreement phenomenon, since ‘when languages mark gender on pronouns and on some other target type, then typically they require a similar machinery to handle them all,’ and since ‘when pronouns are the only evidence for gender, then the resulting gender system seems to be of the same type at that found in some other, fuller system’ (Corbett 1991:169). From this perspective, whether it is exhibited only on pronouns or in a rich-

er array of target types in a language, agreement is a syntactic process. Moreover, if purely inflectional class features are themselves impervious to agreement, as seems to be the case (see Corbett 2006:122-23), then it is no real solution to confine anaphor agreement to operate only with such features. Nevertheless, we will now restrict our attention to cases of phonological agreement that are local and that thus involve ostensibly morpho-syntactic features even on Zwicky's terms.

6.3 Two Apparent PFS Violations in Local Agreement

6.3.1 Godié

The PFS meets a further challenge in the agreement system of the Kru language Godié (also spelled Godie; see Marchese 1986, 1988), which is in many respects similar to Vata and Krahn. Non-human nouns in Godié fall into one of three agreement categories based on natural phonological classes. These are defined by the backness value of the noun-final vowel as shown in (7). As in Vata, Godié exhibits alliterative concord that targets several series of pronouns. Unlike Vata, however, Godié agreement targets also include definite suffixes, demonstratives, and adjectives.

(7)	FRONT → ε	<i>i, ɪ, e, ɛ</i>	<i>mɛ</i> ‘animal’, <i>li</i> ‘spear’
	CENTRAL → <i>a</i>	<i>i, ʉ, ə, ʌ, a</i>	<i>nyidə</i> ‘cooking pot’, <i>suka</i> ‘rice’
	BACK → σ	<i>u, ʊ, o, ɔ</i>	<i>jlu</i> ‘fog’, <i>no</i> ‘alcoholic drink’

As shown in (8a), the definite suffix *-ɛ* and the pronoun *ɛmɛ* agree with the front-vowel-class noun *mɛ* ‘animal’. The phrase ‘big animal’ in (8b) (repeated from Chapter 5 example (11)) demonstrates phonological agreement with the front class. Phonological agreement with the back class in the phrase ‘big man’ in (8b) reflects a redundant generalization, since there is a semantic rule in Godié analogous to rule (5c) for human nouns in Vata.

(8)	a.	<i>mɛ-ɛ</i>	σ	<i>nii</i>	<i>nə</i> ,	<u><i>ɛmɛ</i></u>	<i>plɪ</i>	<i>lɔɛ</i>
		animal-DEF	he	saw	NON-FINAL	it	is	elephant
		‘The animal he saw was an elephant’						
	b.	<i>nyukpɔ</i>	<i>kəd-ɔ</i>	<u><i>nɔ</i></u>	<i>nii</i>	<i>mɛ</i>	<i>kəd-ɛ</i>	
		man	big	this	saw	animal	big	
		‘the big man saw this big animal’						

The productivity of the Godié phonological class-assignment/agreement rules is demonstrated by their use in borrowings. A noun borrowed with a final back vowel, such as *mioko* ‘milk’, takes agreement forms with a back vowel, a noun borrowed with a final front vowel, such as *jakasi* ‘jackass’, takes agreement forms with a front vowel, and so forth.

Since the rules involved in at least definite and adjective agreement are decidedly local, holding over a syntactic domain, Godié agreement must refer to morphosyntactic features. Godié therefore appears to violate the PPFs. There are at least some cases in which agreement refers to an abstract morphosyntactic feature rather than to a phonological element, as demonstrated in (9a), where the noun denotes a human being. Human nouns regularly take back agreement in Godié regardless of the quality of their final vowel, as do certain lexical exceptions such as the noun *pepe* ‘little thing’ in (9b). This tells us nothing about the possibility of phonological agreement in the same language, however, since class systems always have semantic rules that take priority in cases of conflict.

- (9) a. *Dali* ɔ *mu* *Dakpadu*
 Dali he went Dakpadu
 ‘Dali (he) went to Dakpadu’
- b. *pepe* nɔ ɔmɔ *n* *nii*
 little.thing this it I saw
 ‘This little thing, it’s this I saw...’

There is another, more serious consideration, however: like Krahn, Godié agreement involves only three categories. It would therefore in principle be possible to assign nouns to these classes in the lexicon, as in (10) below, just as in Vata. Under such an analysis, even though the Godié class features would be morphosyntactic, at least the syntactic component proper could be kept free of phonological information, which would make for a more restricted grammar, all else being equal.

- (10) a. /XV_[-BK]/ → class 1
 b. /XV_[øBK]/ → class 2
 c. /XV_[+BK]/ → class 3

6.3.2 Abu’

Demonstrative agreement in the Womsis dialect of Abu’ (Nekitel 1986) poses an added twist that makes any lexical interpretation highly unlikely.

Abu' nouns are divided into agreement classes on the basis of their final segments, as in Vata, Krahn, and Godié. Of course, in Abu' it is not final vowels but rather final consonants that define the majority of the noun classes (see Section 1.3.2). As a result, the number of agreement classes is much higher; indeed, the number of morphologically defined noun classes is somewhere in the teens, and as I will discuss in a moment, counting the number of phonologically defined noun classes is not really a meaningful way to understand the system. As in other Arapesh dialects, Abu' agreement is typically alliterative, with agreement targets including verbs and adjectives, i.e., clearly syntactic targets, which means that the features involved in Abu' agreement must be interpreted as morphosyntactic—what Zwicky would call 'local'. Examples of the Abu' agreement pattern are presented in (11). The data comes from Nekitel 1986.

- | | | | | |
|------|----|---|--------------------------------|------------------------------|
| (11) | a. | <i>aleman</i>
man | <i>afu-<u>n</u>eri</i>
good | <i><u>n</u>-ahe'</i>
went |
| | b. | <i>almil</i>
bird | <i>afu-<u>l</u>i</i>
good | <i><u>l</u>-ahe'</i>
went |
| | c. | <i>ihia^hburuh</i>
butterfly | <i>afu-<u>h</u>i</i>
good | <i><u>h</u>-ahe'</i>
went |
| | d. | <i>bahiataf</i>
scaly river fish | <i>afu-<u>f</u>i</i>
good | <i><u>f</u>-ahe'</i>
went |
| | e. | <i>bahiatas</i>
fish.PL | <i>afu-<u>s</u>i</i>
good | <i><u>s</u>-ahe'</i>
went |

As in other Arapesh languages, there are some circumstances in which Abu' agreement is clearly mediated by abstract morphological features. Nouns for male persons like *aleman* 'man' belong to a common class, the pan-Arapesh *n*-class, on the basis of their semantics and, unsurprisingly, regardless of their noun-final segment. We thus find examples of exceptional nouns such as *baah* 'grandfather' in (12a), which would belong to the *h*-class were the noun subject only to the more general phonological agreement rules. Borrowings of the masculine-person category are treated similarly, as shown by the agreement behavior of the borrowed masculine-person noun *katekis* 'catechist' in (12b).

- | | | | | |
|------|----|-----------------------------|--------------------------------|------------------------------|
| (12) | a. | <i>baah</i>
grandfather | <i>afu-<u>n</u>eri</i>
good | <i><u>n</u>-ahe'</i>
went |
| | b. | <i>katekis</i>
catechist | <i>afu-<u>n</u>eri</i>
good | <i><u>n</u>-ahe'</i>
went |

Nevertheless, there is one category of agreeing elements in Abu', demonstratives, which appears to be defined phonologically, and not by way of abstract morphosyntactic features in the lexicon. The reason for this is that demonstrative agreement refers faithfully to the final phonological segment of the noun, and the number of noun-final segments is potentially open-ended due to the continual incorporation of loanwords. According to Nekitel 1986:193,

it would be a bit misleading to go by demonstratives alone in determining the number of noun classes in Abu'. The objection is based on the apparent principle... that the demonstrative reduplicates [the final phonological segment] of the noun it qualifies and it is evident that if a new phonemic sound is introduced... the demonstrative would inflect that sound thus increasing the number of noun classes as instanced by the [borrowed] Tok Pisin word *pater* 'priest'.

Abu' proximal demonstratives are of the form *aCa*; distals are generally of the form *CaCi*. Examples are given in (13). Though the segment *r* is phonemic in Abu', it does not define a native noun class. Yet according to Nekitel, this sound is targeted in demonstrative agreement, as shown in (13a). Example (13b) shows demonstrative agreement with *p*, which is not even a native Abu' phoneme.

- | | | | | |
|------|----|--------------|--------------------------|----------------------------|
| (13) | a. | <i>pater</i> | <i>ara</i> | (no <i>r</i> -final class) |
| | | priest | this | |
| | b. | <i>paip</i> | <i>apa</i> ~ <i>papi</i> | (no native <i>p</i>) |
| | | pipe | this ~ that | |

We can refer to phenomena such as Abu' demonstrative agreement as *literal alliterative concord* because the agreement appears to refer directly to a phonological element as such. The significance of literal alliterative rules is that they cannot be pushed into the lexicon without trivializing the notion of morphosyntactic representations, since to do so would make the set of morphosyntactic features, and thus the key to the theory of inflection, in principle open-ended and limited only by the phonology of loan words.

Abu' demonstratives immediately follow the nouns they modify. It might therefore still be possible to analyze them as clitics on their head nouns. Demonstratives could be treated as underspecified for the quality of their agreeing consonants, with the value provided by spreading from the preceding noun-final segment by a phonological rule. Under such an analysis the demonstrative *mami* 'those' in (14a) below would simply be underspecified for the quality of its consonants at the point at which the syntax sees it. Instead of a syntactic agreement rule targeting demonstratives, the grammar would contain a phonological consonant harmony rule condi-

tioned to apply over the boundary between a noun and a cliticized demonstrative.

- | | | | | | |
|------|----|---------------|--------------------|----------------|--------------------|
| (14) | a. | <i>alemam</i> | <u><i>mami</i></u> | <i>afumi</i> | <u><i>mami</i></u> |
| | | men | those | good | very |
| | b. | <i>naif</i> | <u><i>bief</i></u> | <i>ubahifi</i> | <u><i>fefi</i></u> |
| | | eyes | two | big | very |

There is one test that could be performed to evaluate this proposal. Nekitel notes that the adjective intensifier ‘very’ in Abu’ is identical to the distal demonstrative; indeed, only order distinguishes the two lexemes. Thus, the intensifiers *mami* and *fefi* in (14a) and (14b) must come at the end of the noun phrase, and not immediately following the noun, though in these cases they might still be acquiring their consonantal features by phonological spreading from the syntactically-inflected adjectives that precede them. However, when the head noun is one that ends in a non-native sound and takes abstract agreement by virtue of its masculine-person semantics, like *kiap* ‘patrol officer’, then if the intensifier takes literal alliterative agreement with the noun, it must be by way of a syntactic process, and not by way of a phonological feature-spreading rule, since a differently inflected word intervenes. The hypothesized form is given in (15). Unfortunately, the relevant configuration is not attested.²

- | | | | |
|------|----------------|------------------------|--------------------|
| (15) | <i>ʔkiap</i> | <i>afu-<u>neri</u></i> | <i><u>papi</u></i> |
| | patrol officer | good | very |

Before leaving Abu’, we should also recall that the Rohwim Arapesh pattern of glottal-stop substitution on nouns and corresponding agreement markers discussed in the preceding chapter involves both local and anaphor agreement, in that it carries over from the class III plural noun to all agreement markers in word-final position, regardless of the type of agreement target on which they appear. Thus, *akwokw-i-ga* ‘IV.SG-POSS-III.PL’ agrees anaphorically with the noun *barawa* ‘spears.III.PL’ that occurs in the preceding discourse, and *bia-ga* ‘two-III.PL’ agrees locally in the phrase *na-lib dybariga* ‘bia-ga’ ‘VII.SG-cleared garden.III.PL two-III.PL’. Information about the phonological realization of this noun-final sound must therefore be available to morphosyntax.

²In Hwamsək Abu’, the adjective intensifier is not formally related to the demonstrative; it is realized instead with the invariant form *e’uwe’i*.

6.4 Alliterative Concord in Bainuk

Alliterative concord in the West Atlantic language Bainuk as described by Sauvageot 1967, 1987 is not open to alternative interpretation in terms of phonological harmony.³ Bainuk nouns come in two main types, prefixed and non-prefixed. There are eleven basic classes of prefixed nouns, as represented by the singular-plural pairings in (16).⁴ Examples of prefixed nouns are given in (17). Examples of prefixless nouns are shown in (18).

(16)	<u>Singular</u>	<u>Plural</u>	<u>Other</u>
	1 <i>u-</i>		15 <i>ba-</i>
	2 <i>i-</i>	5 <i>ɲa-</i>	16 <i>di-</i>
	3 <i>ra-</i>		17 <i>ti- ~ pi- ~ bi-</i>
	4 <i>si-</i>	6 <i>mu-</i>	18 <i>ɖa-</i>
	7 <i>gu-</i>	8 <i>ha-</i>	
	9 <i>bu-</i>	10 <i>i-</i>	
	11 <i>ko- ~ kɔ-</i>	12 <i>ko- ~ kɔ-</i>	
	13 <i>da-</i>	14 <i>dĩ-</i>	

(17)	<u>Class</u>	<u>Singular</u>	<u>Plural</u>	
	4 ~ 6	<i>si-nɔx</i>	<i>mu-nɔx</i>	‘tree’
		<i>si-de:n</i>	<i>mu-de:n</i>	‘pirogue’
	7 ~ 8	<i>gu-sɔl</i>	<i>ha-sɔl</i>	‘tunic’
	9 ~ 10	<i>bu-sumɔl</i>	<i>i-sumɔl</i>	‘snake’
		<i>bu-domel</i>	<i>i-domel</i>	‘papaya’

(18)	<u>Singular</u>	<u>Plural (+V[-HI])</u>	
	<i>kata:ma</i>	<i>kata:ma-ã</i>	‘river’
	<i>ɖapɔŋ</i>	<i>ɖapɔŋ-ɔ</i>	‘grass’
	<i>sahri</i>	<i>sahri-ě</i>	‘village’
	<i>ɖiboŋ</i>	<i>ɖiboŋ-õ</i>	‘horse’

Agreement targets include demonstratives, adjectives, interrogatives, pronouns, and the numerals through nine. Thus, as in Godié and Abu’, Bainuk agreement is both anaphoric and local, and the features involved must therefore be morphosyntactic.

³Bainuk (Banyun) is spoken in Senegal and Guinea. Sauvageot 1967 describes the Bignona dialect.

⁴The prefixes in the column labelled ‘Other’ in the chart, numbered fifteen through eighteen, mark the categories diminutive and unlimited vs. limited plurality. I leave them out, somewhat arbitrarily, from my estimation of the number of noun classes.

With prefixed nouns, agreement is alliterative. Bainuk agreement markers match the prefix of the head noun, as shown in (19). The one sort of deviation from this generalization is exemplified by the vowel change in the question *si-nɔx se-rã* ‘which tree?’ in (19a), which seems to be a lowering effect from the following vowel (there is some kind of vowel harmony rule operating in the language). This confirms that, at least with respect to phonological spreading processes, the two words comprise distinct domains. More often, however, agreement markers are identical even phonetically to the prefix of the head noun. The adjectives *si-wuri* ‘long’ and *gu-fēr* ‘white’ in (19e), for example, are inflected in absolute agreement with the prefixes of their respective head nouns, *si-de:n* ‘pirogue’ and *gu-sɔl* ‘tunic’.

(19)	a.	Interrogative	<i>si-nɔx</i>	<i>se-rã</i>
			tree	which
	b.	Numeral	<i>mu-de:n</i>	<i>mu-nak</i>
			pirogues	two
	c.	Demonstratives	<i>si-de:n-o</i>	<i>in-si</i>
			pirogue	this
			<i>ha-sɔl-ɔ</i>	<i>ha-ŋan</i>
			tunics	those (down there)
	d.	Pronouns		<i>in-si</i>
			(pirogue)	this one
				<i>uŋ-gu</i>
			(tunic)	this one
	e.	Adjectives	<i>si-de:n</i>	<i>si-wuri</i>
			pirogue	long
			<i>gu-sɔl</i>	<i>gu-fēr</i>
			tunic	white

Elements agreeing with non-prefixed nouns follow one of two patterns, though it does not seem possible to predict which of them any given unprefixed noun will take. The first pattern involves agreement marking with a constant form: in the singular, the prefix *a-* appears on adjectives and numbers, and the suffix *-no* appears elsewhere. Unprefixed agreement of this sort is shown in (20) below.

- | | | | | | | |
|------|----|----------------|--------------------------|----|--------------|--------------|
| (20) | a. | <i>dibon</i> | <u>a</u> - <i>fēr</i> | b. | <i>dibon</i> | <i>in-no</i> |
| | | horse | white | | horse | this |
| | | <i>sahri-ē</i> | <u>a</u> - <i>wuri-ē</i> | | <i>sahri</i> | <i>in-no</i> |
| | | villages | long | | village | this |

The other pattern of agreement with unprefixing nouns involves an agreement marker in literal alliterative concord with the noun. However, agreement in this case is not with a noun prefix, if the noun has none. Instead, agreement is with the initial CV- sequence of the noun stem. As shown in (21), the unprefixing noun *kata:ma* ‘river’ takes an agreeing demonstrative of the form *in-ka*, the numeral ‘two’ *ka-nak-ā*, the pronominal form *in-ka*, and the adjective ‘large’ *ka-wayi*.⁵ This agreement pattern is neither optional nor peripheral in the language. According to Sauvageot 1987:20, approximately one-third of the 1,200 Bainuk nouns he collected are unprefixing. Sauvageot hypothesizes that many of them are fully assimilated words of foreign origin.⁶

- | | | | | |
|------|----|----------------|------------------|--------------------------|
| (21) | a. | Interrogatives | <i>dapon</i> | <u>d</u> a- <i>rā</i> |
| | | | grass | which? |
| | b. | Numerals | <i>kata:ma-ā</i> | <u>k</u> a- <i>nak-ā</i> |
| | | | river | two |
| | c. | Demonstratives | <i>kata:ma-ŋ</i> | <i>in-<u>k</u>a</i> |
| | | | river | this |
| | | | <i>dapon-ɔ</i> | <i>in-<u>d</u>a</i> |
| | | | grass | this |
| | d. | Pronouns | | <i>in-<u>k</u>a</i> |
| | | | (river) | this one |

⁵The suffix on *ka-nak-ā* is a marker of plurality.

⁶Analyzing another Bainuk variety, Bainouk Gubaher, Cobbinah 2010 finds that the majority of clear borrowings are assimilated to a default class using a constant agreement prefix. He suggests that idiosyncratic alliterating initial syllables may be ‘archaic noun class morphemes in different stages of fusion with the stem’ (189). This is supported for the subset of Gubaher nouns whose initial syllable can be substituted for by prefixes functioning derivationally (e.g., diminutives); it is also suggested by the fact that the idiosyncratically active initial syllables do not exploit the full range of vowel distinctions sanctioned by the language. Cobbinah sees little evidence of literal alliterative agreement at work in Bainouk Gubaher synchronically, though he does consider it possible that the strategy was productive at an earlier stage of the language. As he emphasizes, Bainuk and its relatives are spoken in a situation of intense contact, a characteristic of many other cases where phonological agreement seems to arise. More work on these languages is clearly needed.

e. Adjectives	<i>kata:ma</i>	<i><u>ka</u>-wayi</i>
	river	large
	<i>dapən</i>	<i><u>da</u>-wuri</i>
	grass	long

Again as for Abu', the revealing question is, how many agreement classes are there in Bainuk? In a discussion on how to limit the number of classes posited in an analysis, Dixon proposes that in Bainuk, 'all nouns of the [CV-affixing] type fall into a single... noun class, involving an agreement rule that deals not with a grammatical affix, but instead with the first two phonological segments of a nominal stem' (1982:165-166).

6.5 Implications for Grammatical Organization

The solution Dixon proposes has interesting implications. The categorization of nouns in the CV-affixing class is not determined in the same way that the other classes are, since the information necessary for realizing the CV-affix in agreement with a noun is not given by the morphology of Bainuk. In a sense, the CV-affix is morphologically transparent, much like a skeletally- or prosodically-defined reduplicative affix is transparent to the neighboring phonological material upon which it depends for its realization (Marantz 1982, McCarthy and Prince 1986). Unlike a reduplicative affix, however, which derives its realization from adjacent material within the same word, the realization of the Bainuk CV-agreement affix is dependent on phonological features of an entirely separate word; in other words, the agreement process is syntactic. Put another way, the role of morphology in defining agreement with unprefixing nouns as Sauvageot describes is merely to provide a window of a specified size and placement through which syntax may look to retrieve the appropriate agreement features. What stands behind the window is the phonological form of the noun. Slightly modifying Dixon's account, then, we can propose a general syntactic condition, represented by the noun-adjective agreement rule in (22a). This rule simply requires that nouns and adjectives agree with respect to some feature. Bainuk also has at least twelve specifically morphological agreement classes: eleven basic prefixed categories, as in (22b.i), in addition to the *a-/no* category in (22b.ii). Nouns that fall outside this morphological noun classification system will instead agree phonologically with the head noun, as in (22b.iii).

(22) a. Syntax:

$$N' \rightarrow N_{[\alpha F]} \text{ ADJ}_{[\alpha F]}$$

b. Morphology:

$$(i) \quad X_{[\text{class } 1]} \rightarrow u\text{-}X^\circ \quad (X = N, \text{ ADJ})$$

$$X_{[\text{class } 2]} \rightarrow i\text{-}X^\circ$$

$$\vdots$$

$$X_{[\text{class } 11]}$$

$$(ii) \quad N_{[\text{class } 12]} \rightarrow N^\circ$$

$$\text{ADJ}_{[\text{class } 12]} \rightarrow a\text{-ADJ}^\circ$$

$$\text{ADJ}_{[\text{class } 12, +\text{PL}]} \rightarrow \text{ADJ}^\circ\text{-no}$$

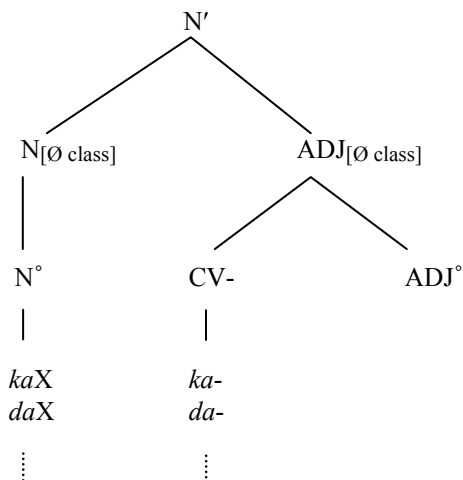
$$(iii) \quad N_{[\emptyset \text{ class}]} \rightarrow N^\circ$$

$$\text{ADJ}_{[\emptyset \text{ class}]} \rightarrow \text{CV-ADJ}^\circ$$

The analytical challenge posed by agreement of this sort is evident when we consider what kind of feature $[\alpha F]$ in rule (22a) refers to. When the head noun belongs to one of the morphologically defined agreement classes, it is reasonable to maintain that agreement involves an abstract feature predicated of the noun, so that $[\alpha F] = \{[\text{class } 1], [\text{class } 2], \dots, [\text{class } 12]\}$. In the case of nouns that fall outside the morphological system, however, the agreement feature appears to be phonological: for these nouns, $[\alpha F] =$ the initial CV of N° , whatever its phonological composition happens to be.⁷ The resulting configuration is expressed in (23), where the morphologically classless adjective is forced to extend one of its branches into phonology directly, unmediated by morphological features, morphosyntactic or otherwise.

⁷Perhaps the rule copies the noun's initial syllable, mora, or series of segments. But whatever the mechanism, real phonological content must be accessed.

(23)



It would of course be possible to treat the literal CV-affixing pattern on a par with the morphological patterns of the regular prefixed noun classes. Under such an analysis each noun of the CV-affixing type would be attributed a unique abstract class feature based on its word-initial phonology, to which a specific morphological realization rule would in turn refer. In that case, however, the set of morphosyntactic features would be suspiciously open-ended, the set of realization rules would exhibit a suspicious redundancy with the phonology of loan words, and the fact that agreement involved sound repetition would be completely coincidental.

In discussing the ordering of attributive adjectives in French in terms of the number of syllables in the adjective and the noun, Miller, Pullum, and Zwicky 1997 face a similar problem: how to express a unitary generalization that must be stated in partly syntactic and partly phonological terms. Before concluding that the relevant facts (*bon vin*, **vin bon* ‘good wine’ vs. *vin ordinaire*, **ordinaire vin* ‘table wine’) are really the result of a usage tendency and are not grammatically driven, Miller, Pullum, and Zwicky 1997:72 ask the following question:

What mechanism could be proposed that would make the information concerning the syllabic structure of the adjective and of the noun available at such phrase nodes [as ADJ' and N'] in order to be able to express the relevant ordering constraints? ...[I]t is quite unclear how we would do it.’

I have not myself developed any satisfactory answer to the same question that the tree in (23) raises, and I share with the authors the feeling that there is something very odd about the structure as it is drawn. But we need

not take the oddness to mean that such a structure could therefore not possibly exist. The case challenges us to think creatively about how to integrate, both conceptually and notationally, qualitatively different types of grammatical information into an analysis.

Bing 1987 represents one effort to do this, suggesting that the phonological agreement pattern exhibited by the Gbobo dialect of Krahn discussed above may be accommodated in the framework of lexical phonology. Although she does not carry through with the analysis, it is not clear this solution is workable, since contrastive information about vowel quality would need to be filled in on Gbobo pronouns post-syntactically, a stage at which the theory only sanctions the manipulation of post-lexical (i.e., allophonic) features. However, I concur with Bing 1987:58 that to convert the phonological information on controller nouns into morphological class features and spreading or unifying them in agreement, only to reconstruct the phonological information at a later stage, ‘disguises rather than explains’ the simultaneously phonological and syntactic nature of the phenomenon.

McLaughlin 1996, 1997 discusses a similar problem in the Niger-Congo language Wolof. Wolof has a small set of classes defined by the initial consonant of the stem, as shown in (24). Interestingly, McLaughlin argues that the Wolof copy rule was used by speakers at an earlier stage of the language specifically as a default strategy, similar to the one described for Arapesh in Chapter 5. Phonological agreement appears to have been triggered when an already-highly alliterative noun class system was confronted with an influx of borrowings on a massive scale (see McLaughlin 1996:444).

(24)	<u>Class</u>	<u>Noun + Determiner</u>	
	G	<i>ginaar gi</i>	‘the chicken’
		<i>gafaka gi</i>	‘the bag’
	S	<i>suukar si</i>	‘the sugar’
		<i>saa si</i>	‘the moment’
	J	<i>jàmm ji</i>	‘the peace’
		<i>jinax ji</i>	‘the mouse’
	W	<i>waxtu wi</i>	‘the hour’
		<i>waasintoor wi</i>	‘the fishscale’

Haude 2006 describes a related situation in Movima, a language isolate spoken in the Bolivian Amazon. Movima has a morphological category of bound elements that are derived by a process of truncation from independent nouns. They are used in compounding, and they are also used anaphorically to cross-reference nouns in discourse—hence their special interest

here. The formal relation of these elements to their source nouns is complex. With native sources, truncated elements are 'homophonous with a particular segment, typically the final syllable, of a corresponding independent noun' (Haude 2006:212): *-ve* < *ju:ve* 'dugout', *-lakwa* < *itila:kwa* 'man', *-mo* < *bove:mo* 'basket'. The pattern deriving truncated elements from Spanish borrowings is more consistent: they are always disyllabic, copying either the last two syllables of three-syllable nouns (e.g., *-su:ka* < *asu:ka* < Spanish *azúcar* 'sugar') or reduplicating the final syllable of two-syllable nouns (e.g., *yaya* < *si:ya* < Spanish *silla* 'chair'). Truncated elements do not serve as shape, consistency, or natural kind classifiers, as do other bound roots in the language. Instead they 'correspond both morphologically and semantically to one particular independent noun' (214); agreement is valid as long as the phonological identity of truncated element to noun-final segment is maintained. But while the truncated elements map to particular nouns formally, the relationship is not fully productive, as there are some loans which cannot truncate or truncate only sporadically. Furthermore, if a conventionalized classificatory bound element exists that could function in place of the truncated form, it will take precedence, just as morphological class agreement in Arapesh takes precedence over purely phonological agreement (so, for example, the native Movima noun *maropa* 'papaya' is not cross-referenced with an element matching its final syllable *pa*, but with the standard classifier used for round-shaped fruits, *ba*:). Again we find direct morphosyntactic manipulation of phonological material that has, 'at some point in time, been a productive process' (215). The regularity with which the process applies to borrowings suggests that it was most active during a time when a situation of intense contact was putting stress on the native morphology to adapt quickly to an influx of new nouns.

There is another case that has been discussed in the literature where phonological agreement operates as a completely regular part of the synchronic language, but the medium of expression is typologically marked. In American Sign Language (ASL), pronouns are formed by pointing to a spatial locus previously established for a nominal referent in the discourse. But the precise locus to which a noun may be assigned is itself unconstrained. Lillo-Martin and Klima 1992 discuss the potentially open-ended range of ASL pronoun forms that result in this situation and conclude, as Dixon does for Bainuk, that there is really only one pronominal element in ASL, and that its realization depends completely on the position in space or 'phonological' form of the syntactically related noun. In a very interesting study exploring the fine variations in the execution of ASL pronouns, Russell and Janzen 2005 argue that their targets are less 'points' than they are regions defined by articulatory targets akin to those speakers aim for in the production of vowels. Thus, despite the descriptive challenges they pose, ASL

pronouns are similar to other phonologically defined categories in that they refer to distinctive regions of the articulatory space, though in the case of sign language that space is manual/visual, as opposed to the articulatory/auditory space out of which phoneme inventories are carved. Leaving aside the question of what sort of features are involved (i.e., whether they are morphosyntactic or purely morphological), we see here another clear instance of a morphosyntactic process extending itself to the limits of a language's phonology.

Phonological agreement is in some ways analogous to the 'repeat classifier' constructions of Southeast Asian languages. In Burmese, for example, there is a subset of classifiers that are either fully or partially identical to the head noun (see Burling 1965, Hla Pe 1965, Lehman 1979, Aikhenvald 2000). Unsurprisingly, these 'echo classifiers' or 'repeaters', which result in constructions such as *qéin ta-qéin* 'house one-house' and *yéitwîn ta-twîn* 'well one-well', have been interpreted as variants of a single phonologically unspecified classifier, with 'the particular allomorph being structurally determined by the form of the antecedent noun' (Burling 1965:250). Lehman 1979:179 n.10 notes that for such classifiers 'the whole lexical entry has *no* phonological part directly attached to it, and one must refer to the entry for [the head noun] for the realization....' For Burmese, many of the nouns that take echo classifiers can also take *-khu*, the default classifier in the language which means something maximally general like 'an instance of' or 'a unit of'. This suggests that repeaters function in a default manner, as do a number of the other cases of phonological agreement discussed here. If repeaters are best characterized as morphosyntactic quantificational devices, as Lehman 1990 argues, then it seems we are dealing with a phenomenon closely related to alliterative concord.

While repeaters can be a source for classifier systems, it is not clear how commonly they serve as a historical source for grammaticized noun class systems. The survey in Aikhenvald 2000 finds no examples of this, though Miraña, a language of Columbia, might be one. Miraña has multiple series of classifiers, many of which are or clearly derive from repeaters. At the same time, these morphemes also function as agreement markers. Seifart 2004:228 describes this as 'a noun class system at an early stage of grammaticalization.' This is interesting because it has been proposed that noun class agreement is derived from repeaters as a synchronic process in Abu' and Bainuk. The approach involves copying the noun in its entirety and moving the copy into position next to partially underspecified agreeing elements, with all but the critical final element of the copy then deleted 'according to some appropriate template' by way of a morphophonological rule (Dimitriadis 1996:9). This is similar to the analysis of Tibeto-Burman classifiers in Lehman 1990, where full-noun copying is also motivated by

surface patterns. The analysis uses only independently attested operations and avoids violation of lexicalism: because the syntactic copy-and-move procedure manipulates the noun only at the level of N° , it need never reach inside it. The idea is to derive agreement by analogy with repeated classifier constructions such as those found in the Daly languages of Australia, in which classifiers are copied on multiple elements in a phrase (see the summary in Corbett 1991:139-141). Literal alliterative concord could then be seen as structurally similar to the incorporated classifier type construction discussed by Rosen 1989, where the same noun is allowed both inside a verb where it contributes to the verb's selectional restrictions, and outside the verb in its usual argument-satisfying position.

The choice between full noun copying plus deletion on the one hand, and overt phonological agreement on the other, cannot really be made on empirical grounds. Instead, we face a metatheoretical trade-off. As Hockett 1947:321 puts it, '[t]he language is not disturbed by our choice; its complexities remain whether itemized in one part or another of our description. But the resulting descriptions may vary a great deal in the clarity with which they depict the situation.' The phonological agreement analysis posits a process that is justified for very few languages, whereas the whole-noun-copy analysis posits processes which are not (for Arapesh, at least) independently justified language-internally. The cases of phonological agreement we are concerned with are quite different from the incorporated classifier construction Dimitriadis takes as a model. In Arapesh, 'incorporation' does not just involve nouns incorporated into verbs, but nouns incorporated into verbs, numerals, adjectives, pronouns, and in most dialects, possessives. And there are no corresponding non-incorporated constructions from which the 'incorporated' elements can be shown to derive.⁸ Moreover, Arapesh nouns are normally completely syntactically acceptable in isolation, which is to say that except for a very few numberless nouns which require the addition of an individuating suffix to be made acceptable as count nouns, nouns do not appear with classifiers. Given the absence of any independent justification for classifiers, it seems that whatever their history, synchronically at least, these systems simply have very concrete agreement.

If the Bainuk and Arapesh phenomena do involve syntactic agreement for phonological features, the dilemma pose is an interesting one. Either syntactic rules must be granted the power to refer to phonological structure

⁸The existence of corresponding non-incorporated constructions is an important tool in tests for classifier incorporation. Indeed, the fact that 'incorporation applies only obligatorily' in certain verbs in West Greenlandic Eskimo leads Rosen 1989:304 n.11 to suggest that the phenomenon should be described as verbal affixation rather than incorporation.

directly, in violation of the PPFS and of lexicalism more generally, or else the interface medium of morphosyntactic representations, and thus the staple of a theory of inflection, must be granted a potential transparency to phonological form that weakens it to the point of superfluity. I do not see how these two alternatives can be decided between on empirical grounds. But in either case, the literal alliterative concordial pattern points to the need for a theory in which grammatical components may interact in a significantly freer way than is generally assumed.

There are several interrelated points to be taken from this discussion; the first concerns methodology. For a proposed principle to be of value in helping us to make analytical choices, it should be possible to state precisely what would constitute a violation of that principle. If literal alliterative concord does not constitute an unequivocal violation of the PPFS, then I think it is fair to say that the PPFS is unfalsifiable on the basis of facts from agreement, because a different interpretation of the data is always available. To that extent, the PPFS makes no predictions about the range of agreement phenomena we should expect to find in the world's languages.

That is not to say that the PPFS makes no predictions about other syntactic phenomena, such as constraints on the ordering of elements in syntactic domains. In a volume devoted specifically to the connection between syntax and phonology, only two out of nineteen contributions even consider the possibility that there is any influence of phonology upon syntax. One of these, Vogel and Kenesei 1990:352, concludes that 'the bulk of the evidence favors a single direction of influence, from syntax to phonology'. The other, Zec and Inkelas 1990, presents evidence that phonology can indeed constrain syntax, but only through units and structures represented at a prosodic level, from which segmental phonology is by definition excluded. Guasti and Nespor 1999 examine the prosodic factors upon constituent order preferences such as the clause-final positioning of long, complex, or 'heavy' NPs. The authors conclude that while relative sonority, number of syllables, and rhythmic considerations like the location of primary word stress may indeed play some role, such prosodic phenomena may only influence stylistic and discourse-based preferences. They never decisively shape syntax proper.

Note that characterizing certain syntactic elements as 'heavy' or 'light' in order to explain their common participation (or non-participation) in syntactic processes is not necessarily to reduce them to phonology. Abeillé and Godard 2001 develop such a weight-based approach to the 'weak' manner and scalar adverbs of French (*bien* 'well', *beaucoup* 'much'), which contrast with other French adverbs in adhering to ordering constraints that generally place them to the left of their complements and constituents they modify. Abeillé and Godard argue that the constraints on the placement of

these adverbs are not easily captured in structural terms; they instead analyze their shared behavior by assigning them the quasi-phonological feature 'lite'. Such a feature is motivated by the adverbs' formal properties, as many (though not all) are monosyllabic and non-derived. However, the feature itself must be an abstract one, because it is not unambiguously predicted by any aspect of phonological or morphological form. Asudeh and Mikkelsen 2000 similarly develop a weight-based analysis of syntactic noun incorporation in Danish, limiting incorporated elements to patients marked as 'lite'; this includes adjectives and common nouns, as well as phrases with lite daughters. Again, the weight-based approach expresses the intuitive notion that the physical composition of syntactic units can affect their distribution, but it does so indirectly by way of abstract features.

It remains to be seen whether we can isolate precise conditions under which phonological defaults and literal agreement can arise. It is surely no coincidence that Bainuk and Arapesh both have large class systems involving a great deal of redundancy in assignment and agreement. Developing the idea that inflectional morphology is based in systems of contrast, Carstairs-McCarthy 1994 speculates that in large overt class systems it is the syntagmatic dependencies themselves that provide the contrastive 'meanings' with which class categories are associated (see also Contini-Morava 2002). While this view leaves unexplained why phonological agreement should be more common in large systems than small ones, it does fit nicely with the suggestion made here that the assignment of class to nouns and the realization of class in agreement may sometimes reduce to two sides of the same phenomenon. On the other hand, Carstairs-McCarthy 1994:781 also predicts that the introduction of new nouns lacking an overt class marker will undermine such a syntagmatic categorization mechanism, encouraging a shift of the burden of contrast away from overt syntagmatic identities and onto covert paradigmatic meanings. But as we have seen, this is not the only possible outcome, as the set of contrasting values may itself be expanded to accommodate the new forms.

Direct acknowledgment of phonological agreement has parallels in the literature on identity in the constraint-based framework of optimality theory, particularly in Yip's (1995, 1998) analysis of identity avoidance in morphology and in discussions of morphological reduplication and truncation. In dealing with patterns of echo-word formation which approach yet veer away from absolute homophony, Yip appeals to a Repeat constraint requiring morphological outputs to contain two identical elements. Partial violations of Repeat are made in deference to a higher-ranked Obligatory Contour Principle that mitigates against adjacent identical elements. This derives the familiar pattern of echo-word formation in which a copy deviates in form, though only minimally, from its base; a classic example is the

Yinglish dismissive 'X-*shm*X' as in *table-shmable*. Within an optimality framework, correspondences among surface forms such as base-truncated form identity and base-reduplicant identity are dimensions along which potential outputs may be evaluated (Benua 1995, McCarthy and Prince 1995). Yip's observation that the calculation of identity is 'fundamentally phonological in nature' (1998:24) leads her to conclude that phonology and morphology may not, after all, comprise discrete grammatical domains.

What the constraint-based analyses share with the cases discussed here is the recognition that morphosyntactic identity can be calculated over what Poplack 1981 calls 'mortal' phonological elements. In the case of phonological agreement, those elements are noun-initial or noun-final sounds. In the Puerto Rican Spanish dialect Poplack studied, the plural-marking *s* phoneme shows a 'concord effect' that is analogous to the glottal substitution found in Arapesh. Poplack found that for plural *s* 'deletion tends to occur from all NP components simultaneously, or is blocked simultaneously on all' (Poplack 1981:70). What we see in all these cases is that 'mortal' sounds whose behavior is firmly rooted in phonology—subject to conditions on word-final deletion or ordinary phonotactic constraints on distribution—are being picked up and used, along with all their idiosyncrasies, by higher levels of grammar to do morphosyntactic work.

Direct reference to the phonological form of a noun in syntactic agreement is eminently reasonable. Indeed, it is perhaps the most efficient and learnable type of agreement rule that one could imagine, since a speaker encountering a single token of a noun will have sufficient information to determine the correct agreement forms, and nothing need be stored in addition to what is in any case necessary for pronouncing the noun itself. We have reviewed evidence that speakers are aware of quite subtle correlations between morphosyntactic status and phonological form in European languages (e.g., Tucker, Lambert, and Rigault 1977; Zubin and Köpcke 1981; and Kelly 1992), and there is certainly no reason to believe that speakers are incapable of making the more concrete correlations involved in alliterative concord. The psychological utility of sound repetition was recognized long ago by Sapir (see also Anttila 1975:24-26), who noted that 'objects... that bear the same imprint are thereby stamped as somehow related' and that although 'sound echo ... [is not] necessary to concord,' such repetition characterizes agreement 'in its most typical and original forms' (1921:114). If syntactic access to phonology were highly unnatural, the phonological similarities that often exist between agreement markers and the elements conveying agreement properties on nouns themselves would be nothing more than inert vestiges of diachrony, and the resounding systematicity of alliterative concord would not be expressible within grammatical theory.

The presence of alliterative concordial patterns has been shown to have a facilitating effect on the learning of phrase structure. When given stimuli embellished with ‘rhyming morphology’, subjects in an artificial language learning situation were able to learn the phrase structure rules of the language more easily than when no such morphology was present, even though it ‘created substantially more complex stimuli than those viewed by No Morphology subjects’ (Morgan, Meier, and Newport 1987:538). This effect was moreover achieved with English-speaking subjects, who would not be familiar with concord cues on the basis of their native language.

Studies of how Bantu languages are acquired by children point to a similar conclusion. In Sesotho the acquisition of class markers on nouns and class marking in agreement each follow regular developmental sequences, and agreement errors are few and not obviously phonologically motivated (the literature on Bantu noun class acquisition is nicely summarized in Demuth and Weschler 2012). The fact that the agreement forms are frequently supported by phonological correlates on noun markers evidently facilitates the acquisition process; it has certainly been shown to do this in other languages (see, e.g., Karmiloff-Smith 1981 on French, Berman 1985 on Hebrew, and Smoczyńska 1985 on Polish). So even where phonological identity later becomes a redundant aspect of agreement in the fully learned system, it serves learners as a useful guide.

In a study that looks (among other things) at children acquiring articles in Italian, Bottari, Cipriani, and Chilosi 1993:361 found that a ‘linear agreement strategy’ is used, whereby ‘a prenominal vocalic segment [i.e., a proto-morpheme that is realized by an unmarked vowel, frequently schwa, at the previous stage] is assigned the same phonetic value as the last vocalic segment of the noun it precedes.... through a sort of coindexing to the last vowel of the Noun’. Thus, Italian children acquire the feminine article *la* earlier than they acquire the masculine *il* despite the fact that nouns of both classes occur with similar frequencies, because *la*, unlike *il*, literally agrees with the final vowel of the noun it modifies. The plural articles *le* and *i* are also learned earlier than the masculine singular article, apparently for the same reason. The study found ‘protracted use of [a] before feminine nouns ending in *a* by [one] child’ who was not, on the basis of other evidence, attempting to phonetically approximate the article directly, but rather ‘coindexing’ his proto-determiner ‘with the final vowel in the noun’ (361-362). Similarly, the authors observed ‘late uses of [ø] in front of masculine nouns requiring *il*,’ where ‘[ø] agrees with the noun ending’, i.e., masculine-final [o] (362). The push for alliterative agreement is so strong that at later developmental stages it can even work backwards, making the noun ending agree with the article rather than the other way around. The example the authors mention (361 n.28) is **le volpe* ‘the.F.PL fox.F.PL’, where the final *i*

of *le volpi* is replaced in a way that can only be interpreted as deference to a phonological agreement principle. What such a situation implies is that the agreement morpheme 'is independent of morphophonemic paradigms insofar as it can be filled with something that satisfies an Agreement relation but that need not necessarily belong to the morphological inventory' (Bottari, Cipriani, and Chilosi 1993:362-363).

The pressure for alliterative concord has also been argued to play a role in the evolution of noun classes. Looking at data from the Nakh-Daghestanian languages, Nichols 1989 notes a correlation between class and noun-initial consonants, and an absence of correlations between class and referential semantics outside of rational nouns. On the basis of these observations and lexical comparisons, Nichols argues that a principle of 'articulatory harmony' may account for the development of class in the language family. In contrast to their cognates in related languages, some Nakh nouns apparently innovated initial consonants which would have been motivated by an impetus for nouns to stand in a relation of articulatory harmony with syntagmatically associated verbal prefixes (just as the child in Bottari, Cipriani, and Chilosi's study temporarily innovated *le volpe*). In other cases, where Nakh nouns appear to bear no word-initial relation to agreement forms, cognate nouns from other Daghestanian languages are found which have retained a harmonic noun-initial consonant. Nichols reads the evidence as pointing to 'a situation in which articulatory harmony was a tendency but not a rule, and in which there was no single mechanism for achieving articulatory harmony' (169). The motive for aligning noun-initial and verb-initial consonants in the first place is hard to know, though Nichols hypothesizes that it evolved when the original function of verb-initial (eventual class agreement) prefixes began to break down, and the semantically based model of classification found for rational nouns was reinterpreted as extending across the entire noun lexicon. Nichols 1989:171 concludes that

[t]here is no clear evidence that Nakh-Daghestanian class was ever based primarily on semantics [in non-human nouns], and no evidence that it ever coded anything.... The phonological form of the root appears to have been the primary determinant of original class classification for non-human nouns, and for these nouns the daughter systems remain semantically arbitrary to this day.

Alliterative concord is psychologically plausible, learnable, and influential in language change. We have now seen several cases where reduplicating a phonological element on another word is used as a strategy for satisfying grammatical agreement, and the consequence for linguistic theory is clear: morphosyntax cannot be categorically denied access to a word's phonological form.

Conclusion

This study of the elaborate noun class systems of the Arapesh languages is really an exploration of the limits human language places on the overtness of grammatical categories. How concrete can morphology be? How far away from meaning can a category lean without relinquishing its morpho-syntactically active status? How superficial can language be in calculating what counts as the ‘same’ morpheme? How deeply into word-form can morphosyntax reach, and how directly can it manipulate the material it finds there? If my efforts have been successful, the answers to these and related questions that have been addressed in the foregoing chapters will be sufficiently interesting and challenging to put the Arapesh languages and sound-based classification on the map for linguistics.

As we have seen, much of the work of Arapesh nominal morphology is performed within lexemes, which are conceived here as non-hierarchical mappings between various types of grammatical information: semantics, morphosyntax, purely morphological features, and phonological form. Paradigmatic associations among singular and plural forms are the basic structures in which the noun classes inhere, and this remains true even in where the singular-plural relationships are insufficiently predictable to relate them through structure building rules. Nevertheless, even the least predictable singular-plural correspondences are typically constrained, as shown by the fact that the morphosyntax commonly fails to recognize them when they exceed certain limits of formal similarity. Redundancy rules or schemas are thus needed in order to express the constraints that hold for these irregular but non-random morphological patterns, and such unifying structures play a central role in organizing the agreement system as well. The class categories defined in this way are organized by motivated extension from a core

form, much as semantic categories are organized around a core meaning. Less expected perhaps is the conclusion that the systematic precedence relations that hold among noun classification principles (semantic > morphological > phonological) do not follow from the Elsewhere condition, a point of logic that governs rule interaction in any domain. Instead, this generalization must be specified as a particularly linguistic fact about the way grammaticized noun class systems are structured.

Arapesh noun class morphology is particularly instructive in the extent to which its categories are rooted in phonological form: the noun-final segments that are predictive of class are not obviously morphemes, i.e., they do not have a representation that is separable from the noun stem, and the number of classes corresponds in a systematic way to the word-final consonant segments defined by the language's phonotactics rather than its morphology. Arapesh shows the place of the phonological classification principle in a typology of noun classification strategies: phonologically-based classification is like the finest-meshed of three big sieves; it holds, normally redundantly, beneath the semantic and morphological classification principles, and its effects only become evident when the other principles for some reason fail. Phonological classification can thus be identified as a type of default strategy that grows out of redundancy, a facilitating resource in categorization.

Finally, the Arapesh languages' unusually concrete noun classes speak to the question of what interactions are possible among grammatical components, specifically, whether syntax may 'see' or make reference to phonological information. Until now, most proposed examples of phonological influence on syntax have been analyzed away, but as the cases surveyed here show, the conclusion that syntax is phonology-free cannot be universally maintained: under certain circumstances syntax may make direct reference to phonological features by matching them in agreement. Whether this takes place through copying, unification, or some other mechanism is immaterial; what matters is that agreement is achieved by matching items that are defined by the way they sound. Either we recognize this phenomenon and accept what it means for grammatical architecture, or else we should admit that the principle of phonology-free syntax does not make an empirically testable claim.

Of course, to identify phonological agreement as possible in natural language is not to say that the phenomenon is common or typical of the languages of the world. Undoubtedly it is a rare phenomenon, one that arises only peripherally even in those systems in which it is best attested. Nevertheless, I do not see why that should require us to categorically restrict the kind of information exchange it involves. Cross-linguistic research has revealed all kinds of minor patterns that were previously thought

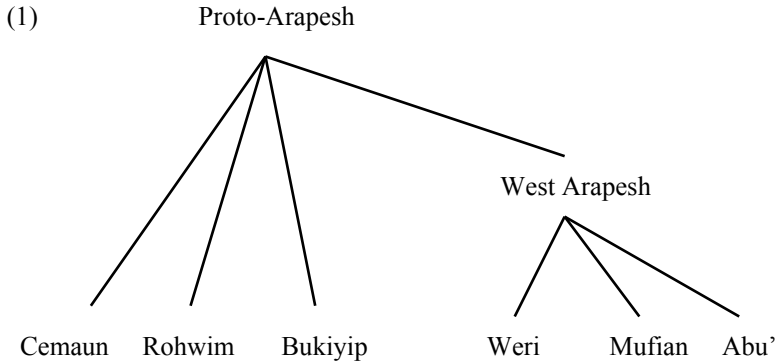
impossible—languages with a preference for onsetless syllables, unlikely morphological categories, object-initial word order, among many other things (Evans and Levinson 2009). I believe we learn more from probing those phenomena that surprise us than we do from reanalyzing them so that they once again look familiar.

Questions certainly remain about how the Arapesh systems arose and transformed over time: What conditions would press languages into developing such phonologically based noun classes? Since phonological classification is so salient and learnable, why do we only find it serving in a default capacity at the bottom end of the typological spectrum, and why is it so rare? Alliteration is a communicatively efficient and commonly attested feature of agreement (Corbett 2006). While sound repetition is clearly one historical source for classifier systems (Aikhenvald 2000, Senft 2000), what would lead an alliterative principle like we find in Arapesh to spread throughout an entire grammaticized system of noun classes or collapse completely?

My best guess is that it represents the extension of a redundant generalization within an already functioning system of semantically and morphologically based classes. Since noun class systems regularly embody redundancy across formal and semantic criteria, a gradual shift of burden to formal criteria could naturally develop if semantic factors became less clear and/or the demand for agreement became stronger, so that classes attracted new members based on their form. If the redundancy became so complete that phonological form was nearly always a sufficient predictor of class, then under situations of duress, such as with a sudden influx of loan words during a new period of language contact, the phonological generalization could become active independently of the other criteria; i.e., it could be elevated to the status of a principle in its own right. That would explain why phonological classification always seems to serve as a last resort, as well as why it tends to occur in large systems: the more redundant categorizations there are, the more specific they will be, the greater confidence speakers will have in their systematicity, and the more likely speakers will be to hypothesize that a unified process underlies them.

There is some indirect evidence that for Arapesh at least, this scenario is correct. Recall that the southwestern-most Arapesh language Weri fails to register a syntactic distinction among sub-classes of non-human nouns; they all constitute a single class for purposes of agreement. The contrast this language provides with its relatives suggests that an interrelated set of morphological conditions support the rationale for phonological classification in those Arapesh languages that exhibit it. Lexicostatistically, Weri is the most divergent member of the Arapesh family. Its highest percentage of cognates with any other Arapesh variety is 57% (Conrad 1978:71). Nekitel

1985 interprets this as indicating that Weri was the earliest language to branch off from proto-Arapesh. However, preliminary evidence from shared innovations justify a western Arapesh subgroup to which Weri belongs (Dobrin 2011):¹



Noun-final inflectional patterns reminiscent of the Arapesh singular-plural pairings are attested elsewhere in the large and diverse Torricelli family, e.g., in Olo (McGregor and McGregor 1982) and One (Donohue 2006). But thoroughgoing phonologically-based classification and agreement with noun-final elements irrespective of semantics was presumably a proto-Arapesh innovation. We can infer from this that Weri must have dropped classification of non-humans after separating from its sisters. What might have motivated this development if so? And why would phonological agreement operate with such crisp regularity throughout the rest of the family if it is archaic?

The paradigmatic opposition of noun-final consonants seems to be the key. The regularity of agreement with subclasses of nonhuman nouns increases as one moves away from Weri and West Arapesh, eastward and especially northward through the family. This distribution is mirrored by a cluster of interrelated morphological properties that support the alignment of singular and plural consonants in noun-final position, in turn supporting their analysis as class markers. Conversely, the more closely one approaches Weri, both geographically and genealogically, the more commonly one finds the following:

¹Diagnostic sound changes in West Arapesh include Proto-Arapesh **ñ*, **f* > *n*, *s* (e.g., Cemaun *ecah^w* vs. Mufian *isaj* ‘net bag’; Bukiyip *culkum* vs. Hwamsək Abu’ *sul’um* ‘fruit, seed’), and prenasalization of voiced stops **b*, **d*, **g*, **g^w* postvocally (e.g., Cemaun *eiguh^w* vs. Mufian *əngof* ‘fish.PL’; Bukiyip *wabil* vs. Mufian *wambel* ‘village’; Cemaun *maduh* vs. Mufian *mandof* ‘vine’).

- Non-contrastive final vowels (sometimes followed by *h*) on nouns (e.g., Weri *usinə* ‘tree kangaroo’ vs. Rohwim *jehin*, Weri *boŋgotəh* ‘tree hollow’ vs. Cemaun *bogot*, Weri *uteŋgeh* ‘cloud’ vs. Cemaun *utəg*). It was the loss of these vowels on singular nouns that led to the morphological contrast between additive and replacive pluralization rules.
- Unpredictable realization of labial release on final *h* in many words, so that the phonemic distinction between *h^w* and *h* is not maintained (e.g., Weri ‘leg’ *bri’ah^w* varies freely with *bri’ah*, ‘basket’ *sirah^w* varies freely with *sirah*).
- A lack of one-to-one correspondence between noun-final consonants in the singular and plural. For example, Weri noun plurals ending in *s* correspond to three distinct plurals in Cemaun, *s*, *c*, and *g^w* (e.g., Weri *roweŋg* ~ *rowəs* ‘tree’ vs. Cemaun *rowog* ~ *rowos*; Weri *witeh* ~ *witous* ‘door’ vs. Cemaun *wit* ~ *witog^w*; Weri *rehin* ~ *rehis* ‘stirred sago’ vs. Cemaun *rehiñ* ~ *rehic*).
- A reduction in the number and variety of word-internal intervocalic change alternations lexically reinforcing the canonical class pairings found noun-finally (e.g., the word-internal alternation *ñ* ~ *c* in Cemaun *amañir* ~ *amaciguh* ‘firefly’ lexically echoing the regular word-final pattern *əriñ* ~ *əric* ‘ember’). In Weri only three canonical pairings are attested in intervocalic changes, and these appear in only a handful of examples; the distribution is similar in Mufian. In Cemaun, by contrast, virtually all the canonical pairings are found noun-internally, and they are extremely frequent. As we have seen, in some environments they even verge on being productive.

The net effect of these patterns is to obscure in Weri the foundational logic that underlies the Arapesh noun classification scheme and preserves it elsewhere in the family: the paradigmatic opposition of noun-final consonants. The class-defining role of noun-final segments is being obscured for similar reasons among the new generation of Cemaun semi-speakers who grew up using Tok Pisin as their first language. Having no word-final *h*, for example—much less a contrast between *h* and *h^w*—these semi-speakers are unprepared to attend to or produce the noun-final phonological contrasts on which the Arapesh system depends. The fact that phonological agreement with subclasses of nonhuman nouns in Arapesh is all or none suggests that there is a grammatical logic inspiring it. In proto-Arapesh that logic was seized upon and elaborated. In Weri it was obscured and dropped.

Perhaps the deepest question this study leaves us with is why there should there exist distinct but overlapping noun classification principles in

the first place. The best answer I have to offer is not my own; it was suggested by Jerry Sadock in 1983. The traditional ideal of efficiency notwithstanding, redundancy is a useful property for languages to have, since they need to suit the limited information processing capabilities of the human beings who use them. But in order for grammatical principles to be redundant, they have to be recognizably distinct; otherwise, they fuse into a single complex principle, and all the benefits of redundancy evaporate. For this reason, we should expect to find slippage among different principles in determining noun class, as we do in determining the order of elements in a string, establishing parts of speech, and conducting analysis in so many other areas of grammar. The governing principles each stand on their own, but they gather their strength from one another.

References

- à Beckett, M. 2010. *Gender Assignment and Word-Final Pronunciation in French: Two Semantic Systems*. LINCOS Studies in French Linguistics 9. Munich: LINCOS.
- Abeillé, A. and D. Godard. 2001. A Class of 'Lite' Adverbs in French. *Romance Syntax, Semantics, and L2 Acquisition: Selected Papers from the 30th Linguistic Symposium on Romance Languages*, ed. J. Camps and C. R. Wiltshire, 9-25. Amsterdam: John Benjamins.
- Adams, K. L. and N. F. Conklin. 1973. Toward a Theory of Natural Classification. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, ed. C. Corum, T. C. Smith-Clark, and A. Wieser, 1-10. Chicago: Chicago Linguistic Society.
- Aikhenvald, A. Y. and R. M. W. Dixon. 1998. Dependencies Between Grammatical Systems. *Language* 74:56-80.
- Aikhenvald, A. Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2004. Language Endangerment in the Sepik Area of Papua New Guinea. *Lectures on Endangered Languages* 5, ed. O. Sakiyama and F. Endo, 97-142. Osaka: Endangered Languages of the Pacific Rim.
- Albright, A. and B. Hayes. 2003. Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study. *Cognition* 90:119-161.
- Allan, K. 1977. Classifiers. *Language* 53:285-311.
- Alungum, J., R. J. Conrad, and J. Lukas. 1978. Some Muhiang Grammatical Notes. *Miscellaneous Papers on Dobu and Arapesh*, 89-130. Workpapers in Papua New Guinea Languages 25. Ukarumpa: Summer Institute of Linguistics.
- Anderson, S. R. 1982. Where's Morphology? *Linguistic Inquiry* 13:571-612.
- Anderson, S. R. 1986. Disjunctive Ordering in Inflectional Morphology. *Natural Language and Linguistic Theory* 4:1-31.

- Anderson, S. R. 1988. Inflection. *Theoretical Morphology: Approaches in Modern Linguistics*, ed. M. Hammond and M. Noonan, 23-43. New York: Academic Press.
- Anderson, S. R. 1992. *A-Morphous Morphology*. Cambridge Studies in Linguistics 62. Cambridge: Cambridge University Press.
- Anttila, R. 1975. *The Indexical Element in Morphology*. Innsbruck: Innsbrücker Beiträge zur Sprachwissenschaft.
- Aristotle. 1954. *Poetics*. Translated by I. Bywater. New York: Random House.
- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry Monograph 1. Cambridge: MIT Press.
- Aronoff, M. 1992. Noun Classes in Arapesh. *Yearbook of Morphology 1991*, ed. G. Booij and J. van Marle, 21-32. Dordrecht: Kluwer.
- Aronoff, M. 1994. *Morphology By Itself*. Linguistic Inquiry Monograph 22. Cambridge: MIT Press.
- Aronoff, M. 1998. Gender Agreement as Morphology. *Proceedings of the First Mediterranean Conference of Morphology*, ed. G. Booij, A. Ralli, and S. Scalise, 7-18. Patras: University of Patras.
- Ashton, E. O. 1944. *Swahili Grammar*. London: Longmans.
- Asudeh, A. and L. H. Mikkelsen. 2000. Incorporation in Danish: Implications for Interfaces. *Grammatical Interfaces in HPSG*, ed. R. Cann, C. Grover, and P. Miller, 1-15. Stanford: CSLI.
- Audring, J. 2009. Gender Assignment and Gender Agreement: Evidence from Pronominal Gender Languages. *Morphology* 18:93-116.
- Bashkow, I. and L. M. Dobrin. 2007. The Historical Study of Ethnographic Fieldwork: Margaret Mead and Reo Fortune among the Mountain Arapesh. *History of Anthropology Newsletter* 34:9-16.
- Battistella, E. L. 1990. *Markedness: The Evaluative Superstructure of Language*. Albany: SUNY Press.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Beard, R. 1995. *Lexeme-Morpheme Base Morphology*. Albany, NY: SUNY Press.
- Becker, A. L. 1975. A Linguistic Image of Nature: The Burmese Numerative Classifier System. *Linguistics* 165:109-121.
- Benua, L. 1995. Identity Effects in Morphological Truncation. *Papers in Optimality Theory*, ed. J. N. Beckman, L. Walsh Dickey, and S. Urbanczyk, 77-136. University of Massachusetts Occasional Papers in Linguistics 18. Amherst: Graduate Linguistics Students Association.
- Berlin, B. 1968. *Tzeltal Numeral Classifiers: A Study in Ethnographic Semantics*. Janua Linguarum Series Practica 70. The Hague: Mouton.
- Berlin, B. and P. Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

- Berman, R. 1985. The Acquisition of Hebrew. *The Crosslinguistic Study of Language Acquisition, Volume 1*, ed. D. I. Slobin, 255–371. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bing, J. M. 1987. Phonologically Conditioned Agreement: Evidence from Krahn. *Current Approaches to African Linguistics* 4, ed. D. Odden, 53-60. Dordrecht: Foris.
- Blevins, J. P. 2006. Word-Based Morphology. *Journal of Linguistics* 42:531-573.
- Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bochner, H. 1993. *Simplicity in Generative Morphology*. Publications in Language Sciences 37. Berlin: Mouton de Gruyter.
- Bokamba, E. G. 1993. Language Variation and Change in Pervasively Multilingual Societies: Bantu Languages. *Topics in African Linguistics*, ed. S. Mufwene and L. Moshi, 207-252. Current Issues in Linguistic Theory 100. Amsterdam: Benjamins.
- Booij, G. 1999. Lexical Storage and Regular Processes. *Behavioral and Brain Sciences* 22:1016.
- Booij, G. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Boroditsky, L., L. A. Schmidt, and W. Phillips. 2003. Sex, Syntax, and Semantics. *Language in Mind: Advances in the Study of Language and Thought*, ed. D. Gentner and S. Goldin-Meadow, 61-80. Cambridge: MIT Press.
- Bottari, P., P. Cipriani, and A. M. Chilosi. 1993. Protosyntactic Devices in the Acquisition of Italian Free Morphology. *Language Acquisition* 3:327-369.
- Boudelaa, S. and M. G. Gaskell. 2002. A Re-examination of the Default System for Arabic Plurals. *Language and Cognitive Processes* 17:321-343.
- Burgmann, A. 1963. [Review of] Gerstners Grammatik der Alubänsprache (Nordküste von Neuguinea). *Anthropos* 58:897-899.
- Burling, R. 1965. How to Choose a Burmese Numeral Classifier. *Context and Meaning in Cultural Anthropology*, ed. M. E. Spiro, 243-264. New York: The Free Press.
- Bybee, J. L. 1988. Morphology as Lexical Organization. *Theoretical Morphology*, ed. M. Hammond and M. Noonan, 119-141. San Diego: Academic Press.
- Bybee, J. L. 1994. A View of Phonology from a Cognitive and Functional Perspective. *Cognitive Linguistics* 5:285-305.
- Bybee, J. L. 1995. Regular Morphology and the Lexicon. *Language and Cognitive Processes* 10:425-455.
- Bybee, J. L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, J. L. and McClelland, J. 2005. Alternatives to the Combinatorial Paradigm of Linguistic Theory Based on Domain General Principles of Human Cognition. *The Linguistic Review* 22:381-410.
- Bybee, J. L. and C. L. Moder. 1983. Morphological Classes as Natural Categories. *Language* 59:251-270.

- Capell, A. 1969. *A Survey of New Guinea Languages*. Australia: Sydney University Press.
- Carstairs-McCarthy, A. 1994. Inflection Classes, Gender, and the Principle of Contrast. *Language* 70:737-788.
- Cassidy, K. W., M. H. Kelly, and L. J. Sharoni. 1999. Inferring Gender from Name Phonology. *Journal of Experimental Psychology: General* 128:362-381.
- Chomsky, N. 1970. Remarks on Nominalization. *Readings in English Transformational Grammar*, ed. R. A. Jacobs and P. S. Rosenbaum, 184-221. Waltham, MA: Ginn.
- Clahsen, H. 1999. Lexical Entries and Rules of Language: A Multidisciplinary Study of German Inflection. *Behavioral and Brain Sciences* 22:991-1060.
- Clements, G. N. 1985. The Geometry of Phonological Features. *Phonology* 2:225-252.
- Cobbinah, A. 2010. The Casamance as an Area of Intense Language Contact: The Case of Bainounk Gubaher. *Journal of Language Contact*, Thema series 3:176-201.
- Conrad, R. J. 1978. A Survey of the Arapesh Language Family of Papua New Guinea. *Miscellaneous Papers on Dobu and Arapesh*, 57-77. Workpapers in Papua New Guinea Languages 25. Ukarumpa: Summer Institute of Linguistics.
- Conrad, R. J. 1987. Kinds of Information in Bukiyip Oral Narrative Discourse. *Language and Linguistics in Melanesia* 16:23-40.
- Conrad, R. J. 1996. More on Arapesh Noun Classes. Unpublished manuscript. Dallas: SIL.
- Conrad, R. J. n.d. Robert Conrad Papers. Mss. 732. Mandeville Special Collections Library, University of California, San Diego.
- Conrad, R. J., and K. Wogiga. 1991. An Outline of Bukiyip Grammar. *Pacific Linguistics* C113. Canberra: Australian National University.
- Contini-Morava, E. 1994. *Noun Classification in Swahili*. Publications of the Institute for Advanced Technology in the Humanities Research Reports, Second Series. <<http://www2.iath.virginia.edu/swahili/swahili.html>>
- Contini-Morava, E. 2002. (What) Do Noun Class Markers Mean? *Signal, Meaning, and Message: Perspectives on Sign-Based Linguistics*, ed. W. Reid, R. Otheguy, and N. Stern, 3-64. *Studies in Functional and Structural Linguistics* 48. Amsterdam: John Benjamins.
- Contini-Morava, E. and M. Kilarski. 2011. Functions of Nominal Classification. Unpublished manuscript. University of Virginia and Adam Mickiewicz University.
- Corbett, G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, G. 2010. Features: Essential Notions. *Features: Perspectives on a Key Notion in Linguistics*, ed. A. Kibort and G. Corbett, 17-36. Oxford: Oxford University Press.

- Craig, C. G., ed. 1986. *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification*. Typological Studies in Language 7. Amsterdam: Benjamins.
- Dahl, Ö. 2000. Animacy and the Notion of Semantic Gender. *Gender in Grammar and Cognition*, ed. B. Unterbeck, 99-115. Berlin: Mouton de Gruyter.
- Dahl, Ö. and M. Koptjevskaja-Tamm. 2001. Kinship in Grammar. *Dimensions of Possession*, ed. I. Baron, M. Herslund, and F. Sørensen. Typological studies in Language 47. Amsterdam: John Benjamins.
- Dahlstrom, A. 1995. Motivation vs. Predictability in Algonquian Gender. *Papers of the Twenty-Sixth Algonquian Conference*, ed. D. H. Pentland, 52-66. Winnipeg: University of Manitoba.
- Demuth, K., N. Faraclas, and L. Marchese. 1986. Niger-Congo Noun Class and Agreement Systems in Language Acquisition and Historical Change. *Noun Classes and Categorization*, ed. C. Craig, 453-471. Amsterdam: John Benjamins.
- Demuth, K. and S. Weschler. 2012. The Acquisition of Sesotho Nominal Agreement. *Morphology* 22:67-88.
- Denny, J. P. 1976. What are Noun Classifiers Good for? *The Proceedings of the Twelfth Regional Meeting of the Chicago Linguistic Society, Vol. 1*, ed. S. Mufwene, C. Walker, and S. Steever, 122-132. Chicago: Chicago Linguistic Society.
- Dimitriadis, A. 1996. *Alliterative Concord in Phonology-Free Syntax*. Unpublished manuscript, University of Pennsylvania.
- DiSciullo, A. M. and E. Williams. 1987. *On the Definition of Word*. Linguistic Inquiry Monograph 14. Cambridge, MA: MIT Press.
- Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 1982. *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Berlin: Mouton.
- Dobrin, L. M. 1997. Lexical Splitting in the Kinship Vocabulary of the Buki (Arapesh) Languages. *Language and Linguistics in Melanesia* 28:99-119.
- Dobrin, L. M. 2001. Arapesh. *Facts About the World's Languages: An Encyclopedia of the World's Languages, Past and Present*, ed. J. Garry and C. Rubino, 33-38. New York: Wilson.
- Dobrin, L. M. 2011. Phonological Classification and Agreement: Evidence from Weri. Poster presented at the Annual Meeting of the Linguistic Society of America, Pittsburgh, PA, January 2011.
- Dobrin, L. M. and I. Bashkow. 2006. 'Pigs for Dance Songs': Reo Fortune's Empathetic Ethnography of the Arapesh Roads. *Histories of Anthropology Annual* vol. 2, ed. R. Darnell and F. Gleach, 123-154. Lincoln, NE: University of Nebraska Press.

- Dobrin, L. M. and I. Bashkow. 2010a. 'Arapesh Warfare': Reo Fortune's Veiled Critique of Margaret Mead's *Sex and Temperament*. *American Anthropologist* 112:370-383.
- Dobrin, L. M. and I. Bashkow. 2010b. 'The Truth in Anthropology Does Not Travel First Class': Reo Fortune's Fateful Encounter with Margaret Mead. *Histories of Anthropology Annual* 6, ed. R. Darnell and F. W. Gleach, 66-128. Lincoln, NE: University of Nebraska Press.
- Donohue, M. 2006. Torricelli Languages. *Encyclopaedia of Language and Linguistics* (2nd ed.), Vol. 13, ed. K. Brown, 1-3. Oxford: Elsevier.
- Dressler, W. U. 1985. On the Predictiveness of Natural Morphology. *Journal of Linguistics* 21:321-337.
- Drosdowski, G., eds. 1984. *Duden Grammatik der deutschen Gegenwartssprache, Band 4*. Mannheim: Dudenverlag.
- Evans, N. and S. C. Levinson. 2009. The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science. *Behavioral and Brain Sciences* 32:429-492.
- Feldman, N. 2005. Learning and Overgeneralization Patterns in a Connectionist Model of the German Plural. Masters thesis, Vienna University.
- Fodor, I. 1959. The Origin of Grammatical Gender. *Lingua* 8:1-41, 186-214.
- Foley, W. A. 1986. *The Papuan Languages of New Guinea*. Cambridge: University Press.
- Foley, W. A. 1991. *The Yimas Language of New Guinea*. Cambridge: University Press.
- Foley, W. A. 2000. The Languages of New Guinea. *Annual Review of Anthropology* 29:357-404.
- Fortune, R. F. 1939. Arapesh Warfare. *American Anthropologist* 41:22-41.
- Fortune, R. F. 1942. *Arapesh*. Publications of the American Ethnological Society 19. New York: J. Augustin.
- Fraser, N. M. and G. G. Corbett. 1997. Defaults in Arapesh. *Lingua* 103:25-57.
- Gerhardt, L. 1994. Western Plateau as a Model for the Development of Benue-Congo Noun-Class Systems. *Afrika und Übersee* 77:161-177.
- Gerstner, P. A. 1939. Der Yams-Anbau im But-Bezirk Neuguineas. *Anthropos* 34: 246-266.
- Gerstner, P. A. 1963. *Grammatik der Alubänsprache*. Micro-Bibliotheca Anthropos, 37. Bonn: St. Augustin.
- Glasgow, D. and R. Loving. 1964. *Languages of the Maprik Subdistrict*. Port Moresby, PNG: Department of Information and Extension Services.
- Goldsmith, J. 2001. Unsupervised Learning of the Morphology of a Natural Language. *Computational Linguistics* 27:153-198.
- Good, J. 2012. How to Become a 'Kwa' Noun. *Morphology* 22:293-335.

- Gray, R., T. Retsema, and R. Hiley. 2008. *A Sociolinguistic Survey of the Abu' Arapesh [aah], Ulau-Suain [svb], and Kap [ykm] Languages*. Dallas: SIL International.
- Greenberg, J. H. 1966. *Language Universals*. Janua Linguarum Series Minor 59. The Hague: Mouton.
- Greenberg, J. H. 1978. How Does a Language Acquire Gender Markers? *Universals of Language*, ed. J. Greenberg, C. A. Ferguson, and E. A. Moravcsik, 73-113. Cambridge, MA: MIT Press.
- Guasti, M. T. and M. Nespor. 1999. Is Syntax Phonology-Free? *Phrasal Phonology*, ed. R. Kager and W. Zonneveld, 73-97. Nijmegen: Nijmegen University Press.
- Harris, J. W. 1991. The Exponence of Gender in Spanish. *Linguistic Inquiry* 22:27-62.
- Harris, Z. S. 1945. Discontinuous Morphemes. *Language* 21:121-127.
- Haspelmath, M. 1989. Schemas in Hausa Plural Formation: Product Orientation and Motivation vs. Source-Orientation and Generation. *Buffalo Working Papers in Linguistics* 89-01:32-74.
- Haspelmath, M. 2006. Against Markedness (and What to Replace It With). *Journal of Linguistics* 42:25-70.
- Haude, K. 2006. A Grammar of Movima. Doctoral dissertation, Nijmegen University.
- Hermann, U. 1985. *Knaurs grosses Wörterbuch der deutschen Sprache*. München: Droemer Knaur.
- Hetzron, R. 1972. Phonology in Syntax. *Journal of Linguistics* 8:251-265.
- Hla Pe, M. 1965. A Re-examination of Burmese Classifiers. *Indo-Pacific Linguistic Studies*, Vol. 2, ed. G. B. Milner and E. J. A. Henderson, 163-185. Amsterdam: North Holland Publishing.
- Hock, H. H. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- Hockett, C. F. 1947. Problems of Morphemic Analysis. *Language* 23:321-343.
- Hockett, C. F. 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- Holton, D., P. Mackridge, and I. Philippaki-Warbuton. 1997. *Greek: A Comprehensive Grammar of the Modern Language*. London: Routledge.
- Hooper, J. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- Jackendoff, R. S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jackendoff, R. S. 1975. Morphological and Semantic Regularities in the Lexicon. *Language* 51:639-671.
- Janda, R. D. 1982. Of Formal Identity and Rule (Un-)Collapsibility: On Lost and Found Generalizations in Morphology. *Proceedings of the First West Coast Conference on Formal Linguistics*, ed. D. P. Flickinger, M. Macken, and N. Wiegand, 179-197. Stanford: Stanford University Linguistics Department.

- Janda, R. D. 1987. On the Motivation for an Evolutionary Typology of Sound-Structural Rules. Doctoral dissertation, University of California, Los Angeles.
- Janda, R. D. 1990. Frequency, Markedness, and Morphological Change: On Predicting the Spread of Noun-Plural -S in Modern High German—and West Germanic. *Proceedings of the Eastern States Conference on Linguistics (ESCOL 90)*, 136-153. Columbus: Ohio State University.
- Janda, R. D. and B. D. Joseph. 1986. One Rule or Many? Sanskrit Reduplication as Fragmented Affixation. *Proceedings of the Eastern States Conference on Linguistics (ESCOL 85)*, 103-119. Columbus: Ohio State University.
- Janda, R. D. and B. D. Joseph. 1992. Meta-Templates and the Underlying (Dis-)Unity of Sanskrit Reduplication. *Proceedings of the Eastern States Conference on Linguistics (ESCOL 91)*, 160-173. Columbus: Ohio State University.
- Janda, R. D. and B. D. Joseph. 1999. The Modern Greek Negator *mi(n)(-)* as a Morphological Constellation. *Greek Linguistics: Proceedings of the 3rd International Conference on Greek Linguistics*, ed. G. Babiniótes and A. Mozer, 341-351. Athens: Elinika Gramata.
- Kapanga, M. T. 1991. Language Variation and Change: A Case Study of Shaba Swahili. Doctoral dissertation, University of Illinois, Urbana-Champaign.
- Karmiloff-Smith, A. 1981. *A Functional Approach to Child Language: A Study of Determiners and Reference*. Cambridge: Cambridge University Press.
- Kay, P. and C. McDaniel. 1978. On the Linguistic Significance of the Meanings of Basic Color Terms. *Language* 54:610-646.
- Kaye, J. D. 1981. La sélection des formes pronominales en Vata. *Revue québécoise de linguistique* 11:117-134.
- Keane, W. 2007. *Christian Moderns: Freedom and Fetish in the Mission Encounter*. Berkeley: University of California Press.
- Kelly, M. H. 1992. Using Sound to Solve Syntactic Problems: The Role of Phonology in Grammatical Category Assignments. *Psychological Review* 99:349-364.
- Kilani-Schoch, M. and W. U. Dressler. 1984. Natural Morphology and Classical vs. Tunisian Arabic. *Wiener Linguistische Gazette* 33/34:51-68.
- Kiparsky, P. 1973. 'Elsewhere' in Phonology. *A Festschrift for Morris Halle*, ed. S. R. Anderson and P. Kiparsky, 93-106. New York: Holt, Rinehart and Winston.
- Kirschbaum, P. Fr., S.V.D. 1921/22. Sprachen- und Kulturgruppierungen in Deutsch-Neuguinea. *Anthropos* 16/17:1052-1053.
- Klaffl, J. and F. Vorman. 1905. Die Sprachen des Berlinhafen-Bezirks in Deutsch Neuguinea. *Mitteilungen des Seminars für Orientalische Sprache* 8:1-138.
- Kolatch, A. J. 1982. *The Name Dictionary: Modern English and Hebrew Names*. Middle Village, NY: Jonathan David.
- Köpcke, K.-M. 1982. *Untersuchungen zum Genusssystem der deutschen Gegenwartssprache*. Linguistische Arbeiten 122. Tübingen: Niemeyer.
- Köpcke, K.-M. 1988. Schemas in German Plural Formation. *Lingua* 74:303-335.

- Koutsoudas, A., G. Sanders, and C. Noll. 1974. The Application of Phonological Rules. *Language* 50:1-28.
- Kulick, D. 1992. *Language Shift and Cultural Reproduction: Socialization, Self, and Syncretism in a Papua New Guinean Village*. Cambridge: Cambridge University Press.
- Lakoff, G. 1986. Classifiers as a Reflection of Mind. *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification*, ed. C. G. Craig, 13-51. Amsterdam: Benjamins.
- Lakoff, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lapointe, S. G. 1988. Toward a Unified Theory of Agreement. *Agreement in Natural Language: Approaches, Theories, Descriptions*, ed. M. Barlow and C. A. Ferguson, 67-88. Stanford: CSLI.
- Lass, R. 1990. How to Do Things with Junk: Exaptation in Language Evolution. *Journal of Linguistics* 26:79-102.
- Laycock, D. 1973. Sepik Languages: Checklist and Preliminary Classification. Pacific Linguistics B25. Canberra: Australian National University.
- Laycock, D. 1975. The Torricelli Phylum. *New Guinea Area Languages and Language Study, 1: Papuan Languages and the New Guinea Linguistic Scene*, ed. S. Wurm, 767-80. Pacific Linguistics C38. Canberra: Australian National University.
- Leavitt, S. C. nd.a. Unpublished Field Notes: Bumbita Arapesh Word List. Mss. 27, Box 3, Folder 5. Mandeville Special Collections Library, University of California, San Diego.
- Leavitt, S. C. nd.b. Unpublished Field Notes: Language Monologues. Mss. 27, Box 3, Folder 6. Mandeville Special Collections Library, University of California, San Diego.
- Leavitt, S. C. nd.c. Unpublished Field Notes: Language Drills. Mss. 27, Box 3, Folder 4. Mandeville Special Collections Library, University of California, San Diego.
- Lehman, F. K. 1979. Aspects of a Formal Theory of Noun Classifiers. *Studies in Language* 3:153-180.
- Lehman, F. K. 1990. Outline of a Formal Syntax of Numerical Expressions, with Especial Reference to the Phenomenon of Numeral Classifiers. *Linguistics in the Tibeto-Burman Area* 13:89-128.
- Lewis, M. P., ed. 2009. *Ethnologue: Languages of the World*. Sixteenth edition. Dallas: SIL International. Online at <http://www.ethnologue.com/>.
- Lieber, R. 1990. *On the Organization of the Lexicon*. Outstanding Dissertations in Linguistics. New York: Garland.
- Lillo-Martin, D. and E. S. Klima. 1992. Pointing out Differences: ASL Pronouns in Syntactic Theory. *Theoretical Issues in Sign Language Research, Volume 1: Linguistics*, ed. S. D. Fischer and P. Siple, 191-210. Chicago: University of Chicago Press.

- López, B. S. 1979. The Sound Pattern of Brazilian Portuguese (Cariocan Dialect). Doctoral dissertation, University of California, Los Angeles.
- Maho, J. 1999. *A Comparative Study of Bantu Noun Classes*. Göteborg: Acta Universitatis Gothoburgensis.
- Marantz, A. 1982. ReReduplication. *Linguistic Inquiry* 13:483-545.
- Marchese, L. 1986. The Pronominal System of Godie. *Pronominal Systems*, ed. U. Wieseemann, 217-255. Tübingen: Gunter Narr.
- Marchese, L. 1988. Noun Classes and Agreement Systems in Kru: A Historical Approach. *Agreement in Natural Language: Approaches, Theories, Descriptions*, ed. M. Barlow and C. Ferguson, 323-341. Stanford: CSLI.
- Marcus, G. F., U. Brinkmann, H. Clahsen, R. Wiese, and S. Pinker. 1995. German Inflection: The Exception That Proves the Rule. *Cognitive Psychology* 41:189-256.
- Matthews, P. H. 1991. *Morphology*. Cambridge: Cambridge University Press.
- Matthews, S. n.d. Material on Language Spoken at Southern (Muhiang) Arapesh Village of Ilahita. Unpublished manuscript obtained by D. Tuzin, 1969.
- Mayerthaler, W. 1981. *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
- McCarthy, J. and A. Prince. 1986. *Prosodic Morphology*. Unpublished manuscript. University of Massachusetts, Amherst and Rutgers University.
- McCarthy, J. and A. Prince. 1993a. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Unpublished manuscript. University of Massachusetts, Amherst and Rutgers University.
- McCarthy, J. and A. Prince. 1993b. *The Emergence of the Unmarked: Optimality in Prosodic Morphology*. Unpublished manuscript. University of Massachusetts, Amherst and Rutgers University.
- McCarthy, J. and A. Prince. 1993c. Generalized Alignment. *Yearbook of Morphology 1993*, ed. G. Booij and J. van Marle, 79-153. Dordrecht: Kluwer.
- McCarthy, J. and A. Prince. 1995. Faithfulness and Reduplicative Identity. *Papers in Optimality Theory*, ed. J. N. Beckman, L. Walsh Dickey, and S. Urbanczyk, 249-384. University of Massachusetts Occasional Papers in Linguistics 18. Amherst, MA: Graduate Linguistics Students Association.
- McClelland, J. and K. Patterson 2002. Rules or Connections in Past Tense Inflections: What Does the Evidence Rule Out? *Trends in Cognitive Sciences* 6:465-472.
- McGregor, D. E. and A. R. F. McGregor. 1982. *Olo Language Materials*. Pacific Linguistics D42. Canberra: Australian National University.
- McLaughlin, F. 1996. Inflection and Phonological Form in Wolof. *1994 Mid-America Linguistics Conference Papers, Volume II*, ed. F. Ingemann, 436-446. Lawrence: University of Kansas Department of Linguistics.
- McLaughlin, F. 1997. Noun Classification in Wolof: When Affixes are not Renewed. *Studies in African Linguistics* 26:1-28.

- Mead, M. 1940. *The Mountain Arapesh: II, Supernaturalism*. Anthropological Papers of the American Museum of Natural History, Volume 37, Part 3. New York.
- Mead, M. 1947. *The Mountain Arapesh: III, Socio-Economic Life, and IV, Diary of Events in Alitua*. Anthropological Papers of the American Museum of Natural History, Volume 40, Part 3. New York.
- Mel'čuk, I. 2006. *Aspects of the Theory of Morphology*, ed. D. Beck. Trends in Linguistics Studies and Monographs 146. Berlin: Mouton de Gruyter
- Mihalic, Rev. F., S.V.D. 1971. *The Jacaranda Dictionary and Grammar of Melanesian Pidgin*. Papua New Guinea: Jacaranda Press.
- Miller, P. H., G. K. Pullum, and A. M. Zwicky. 1997. The Principle of Phonology-Free Syntax: Four Apparent Counterexamples in French. *Journal of Linguistics* 33:67-90.
- Morgan, J. L., R. P. Meier, and E. L. Newport. 1987. Structural Packaging in the Input to Language Learning: Contributions of Prosodic and Morphological Marking of Phrases to the Acquisition of Language. *Cognitive Psychology* 19:498-550.
- Mufwene, S. 1980. Bantu Noun Class Prefixes: Inflectional or Derivational? *Papers from the Sixteenth Regional Meeting of the Chicago Linguistic Society*, ed. J. Kreiman and O. E. Ojedo, 246-258. Chicago: Chicago Linguistic Society.
- Mufwene, S. 1989. la Créolisation en Bantou: Les Cas du Kituba, du Lingala Urbain, et du Swahili du Shaba. *Etudes créoles* 12:74-106.
- Mufwene, S. 1997. Kituba. *Contact Languages: A Wider Perspective*, ed. S. G. Thomason, 173-208. Amsterdam: John Benjamins.
- Nakisa, R. C., K. Plunkett, and U. Hahn. 2002. Single and Dual-Route Models of Inflectional Morphology. *Cognitive Models of Language Acquisition: Inductive and Deductive Approaches*, ed. P. Broeder and J. Murre, 201-219. Oxford: Oxford University Press.
- Nekitel, O. 1984. What Is Happening to Our Vernaculars? *Bikmaus* 2:89-98.
- Nekitel, O. 1985. Sociolinguistic Aspects of Abu', A Papuan Language of the Sepik Area, Papua New Guinea. Doctoral dissertation, Australian National University.
- Nekitel, O. 1986. A Sketch of Nominal Concord in Abu' (An Arapesh Language). *Papers in New Guinea Linguistics No. 24*. Pacific Linguistics A70. Canberra: Australian National University.
- Nekitel, O. 1992. Culture Change, Language Change: The Case of Abu' Arapesh, Sandaun Province, Papua New Guinea. *Culture Change, Language Change—Case Studies from Melanesia*, ed. T. Dutton, 49-58. Pacific Linguistics C120. Canberra: Australian National University.
- Nekitel, O. 1998a. The Semantic Basis of Gender in Abu'. Paper presented at the Annual Meeting of the Linguistic Society of Papua New Guinea. Waigani, Papua New Guinea, July 1998.
- Nekitel, O. 1998b. *Voices of Yesterday, Today and Tomorrow: Language, Culture, and Identity*. New Delhi: UBS Publishers' Distributors.

- Nesset, T. 2006. Gender Meets the Usage-Based Model: Four Principles of Rule Interaction in Gender Assignment. *Lingua* 116:1369-1393.
- Nichols, J. 1986. On Form and Content in Typology. *Language Typology 1985*, ed. W. P. Lehmann, 141-162. Current Issues in Linguistic Theory 47. Amsterdam: John Benjamins.
- Nichols, J. 1988. On Alienable and Inalienable Possession. In *Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, ed. W. Shipley, 557-609. Berlin: Mouton de Gruyter.
- Nichols, J. 1989. The Nakh Evidence for the History of Gender in Nakh-Daghestanian. *The Non-Slavic Languages of the USSR: Linguistic Studies*, ed. H. Aronson, 158-175. Chicago: Chicago Linguistic Society.
- Nidue, J. A. 1990. Language Use in a New Guinea Village: A Triglossic Profile of Makopin I. *Language and Linguistics in Melanesia* 21:47-69.
- Parker, E. M. and R. J. Hayward. 1985. *An Afar-English-French Dictionary (with Grammatical Notes in English)*. London: School of Oriental and African Studies, University of London.
- Paul, H. 1968[1917]. *Deutsche Grammatik, Vol. 2*. Tübingen: Niemeyer.
- Perlmutter, D. 1988. The Split Morphology Hypothesis: Evidence from Yiddish. *Theoretical Morphology: Approaches in Modern Linguistics*, ed. M. Hammond and M. Noonan, 79-100. San Diego, CA: Academic Press.
- Pinker, S. 1999. *Words and Rules*. New York: Basic Books.
- Pinker, S. and A. Prince. 1994. Regular and Irregular Morphology and the Psychological Status of Rules of Grammar. *The Reality of Linguistic Rules*, ed. S. Lima, R. Corrigan, and G. Iverson, 321-351. Amsterdam: Benjamins.
- Pinker, S. and M. T. Ullman. 2002. The Past and Future of the Past Tense. *Trends in Cognitive Sciences* 6:456-463.
- Plaster, K. and M. Polinsky. 2010. Features in Categorization, or a New Look at an Old Problem. *Features: Perspectives on a Key Notion in Linguistics*, ed. A. Kibort and G. G. Corbett, 109-142. Oxford: Oxford University Press.
- Platt, D. 1981. Old Provençal: The Balance Between Regularity and Irregularity in Verbal Inflection. *Inflectional Morphology: Introduction to the Extended Word-and-Paradigm Theory*, ed. T. Thomas-Flinders, 41-71. UCLA Occasional Papers 4.
- Plunkett, K. and R. C. Nakisa. 1997. A Connectionist Model of the Arabic Plural System. *Language and Cognitive Processes* 5/6:807-836.
- Poplack, S. 1981. Mortal Phonemes as Plural Morphemes. *Variation Omnibus*, ed. D. Sankoff and H. Cedergren, 59-71. Alberta: Linguistic Research.
- Prasada, S. and Pinker, S. 1993. Generalization of Regular and Irregular Morphological Patterns. *Language and Cognitive Processes* 8:1-56.
- Pullum, G. K. and A. M. Zwicky. 1988. The Syntax-Phonology Interface. *Linguistics: The Cambridge Survey. Volume 1, Linguistic Theory: Foundations*, ed. F. J. Newmeyer, 255-279. Cambridge: Cambridge University Press.

- Robins, R. H. 1951. *Ancient and Medieval Grammatical Theory in Europe; with Particular Reference to Modern Linguistic Doctrines*. London: Bell.
- Robins, R. H. 1990. *A Short History of Linguistics*. 3rd ed. London: Longman.
- Rosch, E. 1978. Principles of Categorization. *Cognition and Categorization*, ed. Eleanor Rosch and B. B. Lloyd, 27-48. Hillsdale, NJ: Lawrence Erlbaum.
- Rosch, E. and C. Mervis. 1975. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7:573-605.
- Rosen, S. T. 1989. Two Types of Noun Incorporation: A Lexical Analysis. *Language* 65:294-317.
- Ross, J. R. 1970. Metalinguistic Anaphora. *Linguistic Inquiry* 1:273.
- Ross, M. 2005. Pronouns as a Preliminary Diagnostic for Grouping Papuan Languages. *Papuan Pasts: Cultural, Linguistic, and Biological Histories of Papuan-speaking Peoples*, ed. A. Pawley, R. Attenborough, J. Golson, and R. Hide, 15-66. Canberra: Pacific Linguistics.
- Ross, M. n.d. Pronouns as Preliminary Evidence for Grouping Papuan Languages. *Papuan Languages and the Trans New Guinea Phylum*, ed. A. Pawley, M. Ross, and M. Osmond. Canberra: Pacific Linguistics.
- Russell, K. and T. Janzen. 2005. The Categorical Nature of ASL Pronoun Locations. Unpublished manuscript. University of Manitoba.
- Sadock, J. 1983. The Necessary Overlapping of Grammatical Components. *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*, ed. J. F. Richardson, M. Marx, and A. Chukerman, 198-221. Chicago: Chicago Linguistic Society.
- Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and World.
- Sauvageot, S. 1967. Note sur la classification nominale en Bainouk. *La classification nominale dans les langues Nègro-Africaines*, 225-236. Paris: Centre National de la Recherche Scientifique.
- Sauvageot, S. 1987. La linguistique en tant que témoignage historique: le cas du Baynunk. *Contributions à l'histoire du Sénégal*, ed. Jean Boulègue. Cahiers du CRA (Centre de Recherches Africaines), 517-22. Paris: CRA.
- Schmidt, A. 1985. *Young People's Dyrbal: An Example of Language Death from Australia*. Cambridge: Cambridge University Press.
- Schreuder, R. and G. B. Flores D'Arcais. 1989. Psycholinguistic Issues in the Lexical Representation of Meaning. *Lexical Representation and Process*, ed. W. Marslen-Wilson, 409-436. Cambridge, MA: MIT Press.
- Seifart, F. 2004. Nominal Classification in Miraña, a Witotoan Language of Colombia. *Sprachtypologie und Universalienforschung* 57:228-246.
- Senft, G. 2000. What Do We Really Know about Nominal Classification Systems? *Systems of Nominal Classification*, ed. G. Senft, 11-49. Cambridge: Cambridge University Press.

- Serzisko, F. 1982. Gender, Noun Class, and Numeral Classification: A Scale of Classificatory Techniques. *Issues in the Theory of Universal Grammar*, ed. R. Dirven and G. Radden, 95-123. Tübingen Beiträge zur Linguistik 196. Tübingen: Gunter Narr Verlag.
- Siewierska, A. 1998. Nominal and Verbal Person Marking. *Linguistic Typology* 2:1-53.
- Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.
- Silverstein, M. 1976. Hierarchy of Features and Ergativity. *Grammatical Categories in Australian Languages*, ed. R. M. W. Dixon, 112-171. Canberra: Australian Institute of Aboriginal Studies.
- Silverstein, M. 1986. Classifiers, Verb Classifiers, and Verbal Categories. *The Proceedings of the Berkeley Linguistics Society*, 497-514. Berkeley: Berkeley Linguistics Society.
- Silverstein, M. 1987. Cognitive Implications of a Referential Hierarchy. *Social and Functional Approaches to Language and Thought*, ed. M. Hickman, 125-164. Orlando: Academic Press.
- Smoczyńska, M. 1985. The Acquisition of Polish. *Crosslinguistic Study of Language Acquisition, Volume 1*, ed. D. I. Slobin, 595-686. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Spencer, A. 1994. Review of Morphology By Itself. *Language* 70:811-817.
- Spitulnik, D. A. 1987. *Semantic Superstructuring and Infrastructuring: Nominal Class Struggle in ChiBemba*. Studies in African Grammatical Systems 4. Bloomington, IN: Indiana University Linguistics Club.
- Spitulnik, D. 1989. Levels of Semantic Structuring in Bantu Noun Classification. *Current Approaches to African Linguistics, Vol. 5*, ed. P. Newman and R. D. Botre, 207-220. Dordrecht: Foris.
- Stanley, R. 1967. Redundancy Rules in Phonology. *Language* 43:393-436.
- Stump, G. T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Tanenhaus, M. K. and M. M. Lucas. 1987. Context Effects in Lexical Processing. *Spoken Word Recognition*, ed. U. H. Frauenfelder and L. Komisarjevsky Tyler, 213-234. Cambridge: MIT Press.
- Terrill, A. 2002. Systems of Nominal Classification in East Papuan Languages. *Oceanic Linguistics* 41:63-88.
- Thomason, S. G. 1988. Double-Marking in Morphological Change. *Proceedings of the Eastern States Conference on Linguistics (ESCOL 87)*, 296-305. Columbus: Ohio State University.
- T'sou, B. K. 1976. The Structure of Nominal Classifier Systems. *Austroasiatic Studies, Part II*, ed. P. N. Jenner, L. C. Thompson, and S. Starosta, 1215-1247. Oceanic Linguistics Special Publication 13. Honolulu: University Press of Hawaii.

- Tucker, G., W. Lambert, and A. Rigault. 1977. *The French Speaker's Skill with Grammatical Gender: An Example of Rule-Governed Behavior*. The Hague: Mouton.
- Vogel, I. and I. Kenesei. 1990. Syntax and Semantics in Phonology. *The Phonology-Syntax Connection*, ed. S. Inkelas and D. Zec, 339-363. Chicago: University of Chicago Press.
- Wegener, H. 1995. The German Plural and Its Acquisition in the Light of Markedness Theory. *The Development of Morphological Systematicity*, ed. H. Pishwa and K. Maroldt, 245-261. Tübingen: Gunter Narr Verlag.
- Wegener, H. 2002. Aufbau von markierten Pluralklassen im Deutschen: eine Herausforderung für die Markiertheitstheorie. *Folia Linguistica* 36:261-295.
- Welmers, W. 1973. *African Language Structures*. Berkeley: University of California Press.
- Whitely, W. H. 1967. Swahili Nominal Classes and English Loan-Words: A Preliminary Survey. *La Classification nominale dans les langues négro-africaines*, ed. G. Manessy, 157-174. Paris: Centre National de la Recherche Scientifique.
- Wurm, S. A. 1982. Indigenous Languages. *Papua New Guinea Atlas: A Nation in Transition*, ed. D. King and S. Ranck, 34-37. Bathurst, Australia, and Port Moresby, PNG: R. Brown and University of Papua New Guinea.
- Wurm, S. A., D. C. Laycock, and C. L. Voorhoeve. 1975. General Papuan Characteristics. *New Guinea Area Languages and Language Study, Vol. 1: Papuan Languages and the New Guinea Linguistic Scene*, ed. S. A. Wurm, 171-189. Pacific Linguistics C38. Australia: Linguistic Circle of Canberra.
- Wurzel, W. U. 1989. *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer.
- Yip, M. 1995. Repetition and Its Avoidance: The Case of Javanese. *Proceedings of South Western Optimality Theory Workshop 1995*, 238-262. Tucson: University of Arizona Department of Linguistics Coyote Papers.
- Yip, M. 1998. Identity Avoidance in Phonology and Morphology. *Morphology and Its Relation to Phonology and Syntax*, ed. S. G. Lapointe, D. K. Brentari, and P. M. Farrell, 216-246. Stanford: CSLI.
- Zawawi, S. 1979. *Loan Words and their Effect on the Classification of Swahili Nominals*. Leiden: E. J. Brill.
- Zec, D. and S. Inkelas. 1990. Prosodically Constrained Syntax. *The Phonology-Syntax Connection*, ed. S. Inkelas and D. Zec, 365-378. Chicago: University of Chicago Press.
- Zubin, D. and K.-M. Köpcke. 1981. Gender: A Less than Arbitrary Grammatical Category. *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*, ed. R. Hendrick, C. Masek, and M. Miller, 439-49. Chicago: Chicago Linguistic Society.
- Zubin, D. and K.-M. Köpcke. 2009. Gender Control: Lexical or Conceptual? *On Inflection*, ed. P. O. Steinkrüger and M. Krifka, 237-262. Berlin: Mouton de Gruyter.

- Zubin, D. and M. Shimojo. 1993. How 'General' Are General Classifiers? With Special Reference to *ko* and *tsu* in Japanese. *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals*, ed. J. Guenter, B. Kaiser, and C. Zoll, 490-502. Berkeley: Berkeley Linguistics Society.
- Zwicky, A. M. 1984. Autonomous Components and Limited Interfacing: Phonology-Free Syntax, the Halleian Syllogism, and Their Kin. *Papers from the Twentieth Regional Meeting of the Chicago Linguistic Society*, ed. J. Drogo, V. Mishra, and D. Testen, 365-386. Chicago: Chicago Linguistic Society.
- Zwicky, A. M. 1987. *Phonologically Conditioned Agreement and Purely Morphological Features*. Report SRC-87-06 of the Syntax Research Center, University of California, Santa Cruz.
- Zwicky, A. M. 1992. Some Choices in the Theory of Morphology. *Formal Grammar: Theory and Implementation*, ed. R. Levine, 327-371. New York: Oxford University Press.
- Zwicky, A. 1996. Syntax and Phonology. *Concise Encyclopedia of Syntactic Theories*, ed. E. Brown and J. Miller, 300-305. New York: Elsevier Science.
- Zwicky, A. M. and G. K. Pullum. 1983. Phonology in Syntax: The Somali Optional Agreement Rule. *Natural Language and Linguistic Theory* 1:385-402.
- Zwicky, A. M. and G. K. Pullum. 1986a. The Principle of Phonology-Free Syntax: Introductory Remarks. *Interfaces. Ohio State University Working Papers in Linguistics* 32:63-91. Columbus: Ohio State University.
- Zwicky, A. M. and G. K. Pullum. 1986b. Two Spurious Counter-Examples to the Principle of Phonology-Free Syntax. *Interfaces. Ohio State Working Papers in Linguistics* 32:92-99. Columbus: Ohio State University.

Index

- ablaut, 3, 75, 87, 115
Abu', 4, 5, 7, 8 (map), 9, 17-25, 33n,
45, 62, 66, 69, 71-78, 89, 90, 96,
103-104, 117, 141, 156, 164-167,
176, 186
additive vs. replacive rules, 37, 50-
51n, 54-56, 60, 62-64, 72-74, 187
Afar, 16, 101-102
American Sign Language, 175-176
anaphora. *See* agreement,
pronominal
agreement:
alliterative. *See* agreement,
phonological
anaphor. *See* agreement, local
default. *See* noun classification,
default
local, 162-165
phonological, 4, 123-141, 152-
153, 156, 158-182, 184
pronoun, 10, 30n, 32, 33, 125,
159-163, 175-176
as diagnostic of classes, 9-10, 32-
34, 100
Aikhenvald, A., 9n, 82, 95, 105-106,
141, 176, 185
Alungum, J., 7, 23, 62, 126
Anderson, S., 29n, 79, 102, 156-157
Arabic, 116, 159
Arapesh. *See* Rohwim
Arapesh language family, 1-7,
8 (map), 9, 186-187
Arapeshan, 4
Aronoff, M., 2, 12, 13, 27-31, 38, 40,
43-45, 47, 61-63, 66, 97, 102-103,
112, 140
ASL. *See* American Sign Language
assonance. *See* noun classification,
phonological
Audring, J., 141n
Bainuk, 156, 168-173, 176-179
Balif Mufian. *See* Mufian
Bantu languages, 2, 14, 59, 82, 112,
123, 141-153, 181
Beard, R., 9n
Bing, J., 160, 174
Blevins, J., 13, 62
blocking, 52-54, 79, 113, 120-121
Bochner, H., 2, 80
Boikin, 7, 45, 134n,
Bokamba, E., 142-151
Booij, G., 80, 113
borrowings. *See* noun classification,
loan words
Bukiyip, 4, 5, 7, 8 (map), 23, 24, 25,
61-62, 63-64, 66, 69, 89, 96,
100n, 104, 111, 117-119, 126,
128-133, 135, 142, 156, 186
Bumbita. *See* Weri

- Burling, R., 176
 Burmese, 176
 But Arapesh, 66, 69
 Bybee, J., 50, 55, 87-88, 114-116
- Carstairs-McCarthy, A., 13, 70, 160, 179
 classifiers, 11, 40, 66, 81-84, 92, 176-177, 185
 Cemaun, xvi, 5, 7, 23, 24n, 25, 37n, 38, 39-40, 42n, 45n, 47, 56, 66, 67-68, 89-94, 97, 100n, 111-113, 114, 117-120, 126, 134-135, 136n, 145, 186-187
 Clahsen, H., 114
 Clements, G., 114n
 Coastal Arapesh, 8 (map). *See also* Bukiyip; Cemaun; Rohwim
 Cobbinah, A., 170n
 cognates, 7, 186-187
 concord. *See* agreement
 concreteness, xi, 1, 123, 140, 177, 180, 183, 184
 Conrad, R., 4, 5, 7, 23, 39-40, 61-62, 89n, 126, 128, 132n, 135, 185
 contact. *See* language contact
 Contini-Morava, E., 12, 15, 66n, 82, 179
 Corbett, G., 9-10, 14n, 15-16, 30n, 43n, 53, 64-65, 66n, 95, 98, 101, 105n, 108, 113, 121-123, 141, 155, 160, 162-163, 177, 185
 Craig, C., 82
- Dahl, Ö., 10n, 44n
 Dahlstrom, A., 64
 default. *See* agreement, phonological; noun classification, default; noun classification, minority default
 Demuth, K., 60, 101, 181
 Denny, J., 66n, 81, 94
 derivational affix, 11, 75, 97
 diminutives, 20, 25, 61, 66, 75, 107-108
 Dimitriadis, A., 176-177
 Dixon, R., 10, 66n, 82, 84, 105, 123, 171, 175
- Dobrin, L., 4, 5, 6n, 44n, 106, 186
 Donohue, M., 4
 double marking. *See* intervocalic change
 Dressler, W., 55
 Dyrbal, 10, 82-4, 86-8
- ehas* plural, 36, 42-44, 47-54, 56-57, 79, 111-122, 127, 129-130, 134n, 138-140. *See also* noun classification, default
 Elsewhere condition, 25, 47, 102-109, 184
 English, xii, 3, 9-10, 11-12, 13, 15, 28-29, 48, 49n, 50, 67-69, 84, 87-88, 94, 95, 115-116, 125, 148, 181
 Evans, N., 185
 exceptions. *See* irregularity
- Foley, W., 3-4, 9, 23, 102
 Fortune, R., xv, 5, 6, 25, 27, 32n, 33, 37-43, 45, 47, 48, 49-51, 56n, 60-61, 65, 70, 84-85, 89-90, 93, 123, 126-127, 129-130, 133-136
 Fox, 64-65
 Fraser, N., 53-54, 113
 French, 2, 10-12, 14, 173, 178-179
 frequency, 3, 47, 84-85, 89-90, 114-117
- gender. *See* noun classification
 German, 2, 6, 10-14, 47, 55n, 87-88, 107-108, 112-113, 116, 120
 Gerstner, A., 5, 6, 61, 130n
 Godié, 16-17, 123-124, 163-164
 Good, J., 152
 Greek, 10-11, 22n, 108
- Harris, J., 60
 Harris, Z., 70n
 Haspelmath, M., 116
 Haude, K., 174-175
 Hetzron, R., 158
 Hockett, C., 177
 human nouns. *See* kinship terms
 Hwamsək Abu'. *See* Abu'

- Ilahita, 7, 63, 89, 186n. *See also*
 Southern Arapesh
 incorporation, 177
 inflectional classes. *See* realization
 classes
 intervocalic change, 38, 48-49, 74,
 88-94, 187
 irregularity, 3, 4, 10-11, 71-80, 84-
 95, 111, 114-121, 129, 158-159
 Italian, 181-182
- Jackendoff, R., 79, 156
 Janda, R., 12, 47, 50, 69, 77, 80, 88,
 94, 102, 113
 Joseph, B., 12, 50, 77, 80, 88
- Kapanga, M., 142, 145, 150n, 151,
 152
 Keane, W., x-xi
 Kelly, M., 11-12, 180
 KiKongo, 142-143, 145-146
 Kilani-Schoch, M., 55
 Kinshasa LiNgala, 112, 142-144,
 146-148, 151
 kinship terms, 35-36, 40-41, 43-44n,
 56n, 65, 92-93, 96-97, 102-104,
 106-107
 Kiparsky, P., 102
 Kirschbaum, F., 59-60, 62
 KiSwahili. *See* Swahili
 KiTuba, 112, 142-144, 146, 148
 Köpcke, K.-M., 11, 50, 87-88, 108n,
 120
 Krahn, 160, 164, 165, 174
 Kru languages, 16, 123, 152-153.
See also Godié; Krahn; Vata
 Kulick, D., 9
- Lakoff, G., 10, 82-83, 87-88, 94
 language contact, 170n, 175, 185
 language shift, xii, 9
 Lass, R., 12
 Latin, 10, 13, 29-30, 62, 69-70, 98
 Laycock, D., 4, 23, 59-60
 Leavitt, S., 9, 98-99, 100n
 Lehman, F., 176
 Levinson, S., 185
- lexical listing. *See* irregularity;
 schema
 lexicalist hypothesis, 3, 155-158,
 178. *See also* phonology-free
 syntax
 LiNgala, 142-144, 146-148
 literal alliterative concord. *See*
 agreement, phonological
 loans. *See* noun classification, loan
 words
 local agreement. *See* agreement,
 local
- Maho, J., 2
 Marchese, L., 16, 61, 123, 163
 Marcus, G., 114, 116
 markedness, 47-56, 184-185
 Matthews, P., 28
 Matthews, S., 5, 7, 61, 63
 McClelland, J., 114
 McLaughlin, F., 174
 Mead, M., 49n
 meta-rule. *See* schema
 meta-template. *See* schema
 Miller, P., 3, 155, 173
 minority default, 116
 Miraña, 176
 morphological status of Arapesh
 terminals. *See* suffixal status of
 noun class terminals
 morphomic, 13, 28
 morphotactic transparency, 55-56,
 72
 Mountain Arapesh, ix-xi, 12. *See*
also Bukiyip; Cemaun; Rohwim
 Movima, 174-175
 Mufian, 4, 5, 7, 8 (map), 23, 24, 25,
 61-63, 69, 104, 106, 111, 117-
 118, 126, 128-133, 142, 156, 186-
 187. *See also* Ilahita
 Mufwene, S., 66n, 142
- Nakh-Daghestanian, 182
 Nekitel, O., 4-5, 7, 9, 17-24, 62, 71-
 72, 75n, 78n, 102n, 164-167, 185-
 186
 Nessel, T., 102n
 Nichols, J., xi, 1, 12, 182

- Nidue, J., 9
 noun class suffixes. *See* suffixal
 status of noun class terminals
 noun classification:
 arbitrariness of, x, 10
 cultural basis of, 10-11, 83
 default, 3, 20-21, 40-54, 56-57,
 111-132, 174, 185. *See also*
-ehas plural
 definition of, 9-10, 123
 derivational use of markers, 65-
 67, 92
 form-based assignment, 11-12,
 15-16, 121-122. *See also* noun
 classification, morphological;
 noun classification,
 phonological
 loan words, 15, 21, 24n, 39-40,
 45, 49-50, 54-56, 84, 100, 104,
 112, 116-120, 126-132, 134-
 135, 137-141, 152, 165-167,
 170, 173, 174-175
 morphological assignment, 3, 11,
 14-16, 20, 77, 121-122
 number of classes. *See* noun
 classification, size of system
 phonological assignment, ix-xi, 1-
 2, 3, 4, 11-12, 16-25, 83-87
 relations among principles, 3, 20,
 25, 31, 81, 97-98, 101-109,
 111-112, 121-122, 137-141,
 184
 semantic assignment, ix-xi, 2, 10-
 11, 19, 20, 38-40, 77-78, 81-
 84, 96-98, 100-102, 128-129,
 131
 size of system, 2, 6-7, 14, 16, 33,
 165, 171
- Perlmutter, D., 75, 76, 79
 person, 6, 33
 phoneme inventory, 17 (Abu'), 133-
 136 (Rohwim)
 phonology-free syntax, 3, 155-163,
 177-178, 182, 184
 phonotactic constraints, 20, 22-25,
 133, 141
 Pinker, S., 114-115
 Plaster, K., 83
 Plunkett, K., 116
 plural assignment. *See* realization
 classes
 Polinsky, M., 83
 Poplack, S., 180
 Principle of Phonology-Free Syntax.
See phonology-free syntax
 productivity, 39-40, 47, 49-50, 54,
 84, 115-121. *See also*
 markedness; noun classification,
 loan words
 Pullum, G., 3, 155, 157-159, 173
- radial categories, 82-88, 93, 94
 realization classes, 13, 29-31, 32-33,
 37-38, 42-47, 52-53, 112
 redundancy, 184, 185, 188
 redundancy rules. *See* schemas
 redundant marking. *See* intervocalic
 change
 reflexive coding, 12-14
 repeaters, 176-177
 replacive rules. *See* additive vs.
 replacive rules
 Rohwim, xv, 2, 4-6, 21, 24, 25, 27,
 31-57, 64-66, 69, 74-75, 84-86,
 90-91, 93, 96, 100n, 102, 113,
 122-141, 142, 167, 186-187
 Ross, M., 4, 100n
 rule constellations. *See* schemas
- Sadock, G., 98, 188
 Sanskrit, 22n, 88
 Sapir, E., 180
 Saussure, F. de, x
 Saussurean dogma, xi, 1, 12
 Sauvageot, S., 168-171
 schemas, 12-13, 52-53, 77-80, 81-95,
 113-114, 120-121, 183
 Seifart, F., 176
 semiotic ideology, x-xi
 Senft, G., 185
 Shaba Swahili, 112, 142, 144-145,
 148-151
 Silverstein, M., 82, 105-106, 123
 Somali, 158-159

- sound system. *See* phoneme
 inventory
 Southern Arapesh, 4, 5. *See also*
 Ilahita; Mufian
 Spanish, 10, 60, 62, 175, 180
 Spencer, A., 43n
 Spitulnik, D., 82
 suffixal status of noun class
 terminals, 60-71
 Swahili, 12, 14-16, 66n, 70n, 82,
 108-109, 123-125, 141-142, 144-
 145, 148-152. *See also* Shaba
 Swahili
 Standard Swahili. *See* Swahili

 Terrill, A., 4
 Thomason, S., 89n
 Tok Pisin, xii, 9, 17, 22, 24, 39-40,
 49-50, 54-56, 72, 75n, 89n, 90,
 104, 112, 117-120, 126-132, 135,
 142, 145, 166, 187

 Torricelli phylum, 4, 100, 186

 unpredictability. *See* irregularity

 Vata, 159-164

 Wautogik Arapesh. *See* Cemaun
 Wegener, H., 55n
 Weru, 4-5, 7, 8 (map), 83, 89, 92-93,
 96, 98-100, 106, 126n, 185-187
 Wolof, 174
 Womsis Abu'. *See* Abu'
 Wurm, S., 3, 23

 Yiddish, 75-77, 79
 Yimas, 102-103
 Yip, M., 179-180


 Zawawi, S., 152
 Zubin, D., 11, 40, 108n
 Zwicky, A., 3, 155-163, 165, 173

Based on an exhaustive search of published sources and the author's firsthand fieldwork, *Concreteness in Grammar* explores the role of phonological form in the noun class systems of the Arapesh languages spoken in Papua New Guinea. Linguists have long known that form plays a critical role alongside semantics in the classification of lexical items. In Arapesh, virtually every possible final ending of a noun is represented in the paradigm of noun class and agreement markers, reflecting an interpenetration of sound structure and grammar that some theories would disallow as wildly unconstrained. In this book, Lise Dobrin describes these formal patterns in order to reveal their naturalness and elegance, establishing their place in a typology of noun class systems and drawing out their significance for theories of grammatical architecture.

A rigorous study of a group of endangered languages, *Concreteness in Grammar* revisits the definition of the morpheme and looks at unusual language patterns to reveal the naturalness of sound as an organizing feature of grammar.

Lise M. Dobrin is Assistant Professor of Anthropology and Director of the Linguistics Program at the University of Virginia.

Stanford Studies in Morphology and the Lexicon

 **CSLI**
PUBLICATIONS
Center for the Study of
Language and Information
Stanford, California

ISBN 13 978-1-57586-606-2

ISBN 10 1-57586-606-4



9 781575 866062