

The Pragmatics of Word Order

Empirical Approaches to Language Typology

7

Editors

Georg Bossong
Bernard Comrie

Mouton de Gruyter
Berlin · New York

The Pragmatics of Word Order

Typological Dimensions of Verb Initial Languages

Doris L. Payne

Mouton de Gruyter
Berlin · New York 1990

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter & Co., Berlin.

- ⊗ Printed on acid free paper.
(ageing resistant — pH: 7, neutral)

Library of Congress Cataloging in Publication Data

Payne, Doris L., 1952-

The pragmatics of word order : typological dimensions of verb
initial languages / by Doris L. Payne.

p. cm. — (Empirical approaches to language typology ; 7)

Includes bibliographical references (p.) and index.

ISBN 0-89925-612-0 (alk. paper)

1. Grammar, Comparative and general — Word order. 2. Typo-
logy (Linguistics) 3. Pragmatics. I. Title. II. Series.

P295.P35 1990

415 — dc20

90-6041

CIP

Deutsche Bibliothek Cataloging in Publication Data

Payne, Doris L.:

The pragmatics of word order : typological dimensions of verb
initial languages / by Doris L. Payne. — Berlin ; New York :
Mouton de Gruyter, 1990

(Empirical approaches to language typology ; 7)

ISBN 3-11-012207-3

NE: GT

© Copyright 1990 by Walter de Gruyter & Co., D-1000 Berlin 30.

All rights reserved, including those of translation into foreign languages. No part of this
book may be reproduced or transmitted in any form or by any means, electronic or
mechanical, including photocopy, recording, or any information storage and retrieval
system, without permission in writing from the publisher.

Printing: Ratzlow, Berlin. — Binding: Dieter Mikolai, Berlin. — Printed in Germany.

*To the memory of Lucía Macedo,
an eternally great woman,
and to all those she represents*

Acknowledgements

Rarely is any significant work accomplished by one person in isolation. This is not to imply that the following pages are necessarily that significant, but it is true that they have not been done by one person, nor in isolation. This study owes its existence primarily to Des Derbyshire, who I am honored to consider a friend and mentor. He came through town one purported Peruvian dry season and perhaps inadvertently convinced me that one ought to know the basic constituent orders of the language one is working on. I decided to take an afternoon off from looking at phonology to definitively settle the question for Yagua. This, and much more, is the result.

Pedro Díaz, Gloria Cahuachi de Díaz, and other members of the community of Urco Miraño not only helped Tom Payne and myself learn something about Yagua, but also offered their friendship and put up with our lack of social graces. By sharing their home, Alchico Jápiirýa and Estela Múcatyurirýa taught us a great deal, especially about the unimportance of manufactured goods. Pedro Díaz, Mamerto Macahuachi, Hilario Peña, and Alcides Lozano Salazar gave unselfishly of their time, energy, and patience. Paul Powlison shared his years of knowledge about jungle living and the Yagua language, culture, and an excellent text collection, as well as many helpful comments along the way. Each of these people has enriched our lives and this study would not have been possible without them. Various members of the Summer Institute of Linguistics provided technical support in the form of air and river transportation, computer technology and software, and the underrated art of babysitting. The Peruvian Ministry of Education made it possible for me to pursue this research in Peru. This work was partially supported by the UCLA Graduate Division and the Department of Linguistics. I thank each of these people and institutions for their support.

Many thanks are also due to Pamela Munro, Jack Du Bois, Ed Keenan, Paul Kroskrity, Carlos Quicoli, Sandy Thompson, Tom Payne, Steve Anderson, Bernard Comrie, and Marie-Louise Liebe-Harkort for helpful comments at various points. Donna Fisher and Tom Payne professionally prepared the manuscript.

The people who facilitated this research in the most consequential way are those who shared in the lives of my daughters Claire and Stephanie during its duration: Delicia Méndez, Kimber Olson-Brady, Chela Díaz, Margaret and Russ Obenchain, Barbara Hagan, Suhair Azzam, Shirean Williams, Jan Tyhurst, Taís and Richard Weisenberger, Sheila Fountain,

viii *Acknowledgements*

the staff of the UCLA Parent-Toddler Co-op, and the staff of the Marina Christian Preschool. Finally, my most heartfelt thanks go to my family. During the most solid period of writing Anna Claire taught me that playing fairies is the most important thing in life and that linguists are not scientists; Stephanie Joy was the delight of my life simply by virtue of being two. And always, grateful thanks to Tom. If there is anything of value in the following pages, either in terms of perceived fact or proposed explanation, *Soli Deo Gloria*.

Eugene, Oregon
January 1990

Contents

Acknowledgements	vii
Abbreviations	xiii
 Chapter One	
Introduction	1
1.1. Genetic and typological affiliations	2
1.2. Demography and ethnography	4
1.3. Previous linguistic work on Peba-Yaguan	5
1.4. Data for the current study	6
 Chapter Two	
Constituent Order and Order Correlations	9
2.1. Observations of constituent order co-occurrences	9
2.2. The verb initial norm (VIN)	10
2.3. Selected theoretical approaches accounting for word order correspondences	16
2.4. Identification of basic constituent order	21
2.5. Towards an adequate constituent order typology	23
 Chapter Three	
Clausal Phenomena	27
3.1. Major structural clause types	26
3.1.1. Clause Type 1	29
3.1.2. Clause Type 2: S_0 clauses	38
3.1.3. Clause Type 3: Predicate nominals and predicate locatives	39
3.1.4. Type 1 predicate nominals	40
3.2. Impersonals and functionally related constructions	41
3.2.1. The impersonal construction	42
3.2.2. The anti-causative	43
3.2.3. Predicate nominals with object nominalizations	44
3.2.4. Lexical passives	45
3.3. Auxiliaries	45
3.4. Second position clitics	48
3.4.1. Second position clitics within C^2	48
3.4.2. Second position clitics in C^1	53
3.4.3. Constituency of auxiliary plus main verb	58
3.5. Causation and desideration	60

3.6. Parataxis	62
3.7. Negatives and modals	64
3.7.1. Negatives	64
3.7.2. Modals	68
3.8. Questions	69
3.8.1. Yes-no questions	69
3.8.2. Information questions	70
3.9. Comparatives and equatives	73
3.10. Coordination and alternative relations	74
3.11. Complex sentences	76
3.11.1. Unmarked sentential complements	77
3.11.2. Marked sentential complements	78
3.11.3. Adverbial clauses with <i>-ñy</i> and other conjunctions	80
3.11.4. Relative clauses	82
3.11.5. Correlative clauses	90
3.11.6. Indirect quote complements	92
3.11.7. Infinitival adverbials	92
3.11.8. Infinitival complements and verb serialization	94
3.12. Summary	95
 Chapter Four	
Noun and Postpositional Phrase Phenomena	97
4.1. Bound modifying roots	99
4.2. Determination of head versus modifier within noun phrases	102
4.2.1. Category constancy	103
4.2.2. Unique immediate constituent, and obligatorily present	103
4.2.3. Subcategorization and government	104
4.2.4. Pragmatic head	106
4.3. Order of head noun and descriptive modifier in text	111
4.4. Complex modifying phrases	119
4.5. Genitives	119
4.6. Postpositional phrases	122
4.7. Summary	127
 Chapter Five	
Noun Classification and Nominalization	129
5.1. Derivational uses of classifiers	130
5.2. Inflectional uses of classifiers	134
5.3. Anaphora and classifiers	140
5.4. Theoretical status of Yagua classifiers	143
5.4.1. Analysis I	146

5.4.2. Analysis II	146
5.4.3. Analysis III	148
5.4.4. Inflection versus derivation within a prototype framework	150
Chapter Six	
The Verb Phrase and Related Issues	159
6.1. Verbal nexus	159
6.1.1. Same-subject infinitival complements	160
6.1.2. Verb serialization	164
6.2. Adverbs	166
6.3. Subject - object asymmetries: Evidence for a verb phrase containing the object?	167
6.4. Incorporation	172
6.5. An overview of verbal morphology	172
6.6. The instrumental/comitative <i>-ta</i>	174
6.7. Morphological causatives with <i>-tániy</i>	180
6.7.1. Morphology of the causative verb	180
6.7.2. Set II reference and order of arguments with <i>-tániy</i> causative	184
6.8. Summary	188
Chapter Seven	
Pragmatic Factors Motivating Order Variation	189
7.1. General pragmatic structure of Yagua clauses	192
7.2. The pragmatically marked nucleus	195
7.3. Function of the PM' component	197
7.4. Functions of the PM component	197
7.4.1. Single focus contrast and other single focus subtypes	198
7.4.2. Multiple foci of contrast	202
7.4.3. Questions and answers to information questions	204
7.4.4. Restatement and added detail restatement	207
7.4.5. Counter expectation	212
7.4.6. Threats	213
7.4.7. Negation	215
7.4.8. Semantically marked conditions	215
7.4.9. Problem cases	216
7.5. Summary of pragmatically marked types	218
7.6. Frequency distribution of syntactic constituent orders	222

xii *Contents*

7.7. Relative order of direct objects and obliques	230
7.8. Summary	237
Chapter Eight	
Constituent Order in Yagua: Conclusions and Implications	239
8.1. Arguments in favor of SVO as basic	239
8.1.1. Hawkins' Universal 2	239
8.1.2. Set I clitic reference	240
8.1.3. Subject - object asymmetries	243
8.2. Arguments against SVO as basic	243
8.3. Summary of typological traits	246
8.4. Implications for head-dependent ordering principles and Hawkins' Universals	248
8.5. Yagua as a head marking language	250
Appendix 1: Lagarto (Alligator) Text	257
Appendix 2: Phonology	265
Notes	269
References	281
Index	291

List of Abbreviations

A	subject of a transitive clause	JUU	modal clitic <i>-juu</i>
AB	ablative	LOC	locative
ACHIEVE	action achieved (meaning is actually uncertain)	MALF	malefactive
AL	allative	NEG	negative
ANIM	animate	NEUT	neutral classifier (animate or inanimate)
ANTCAUS	anticausative	NIITA	second position clitic <i>-niita</i>
CAH	Cahocuma dialect	NIY	second position clitic <i>-niy</i> 'contrast'
CAUS	causative	NMLZR	nominalizer
CAY	modal clitic <i>-cay</i>	O	object; occasionally object or oblique
CO	modal clitic <i>-cô</i>	O:NOM	nominalizer on understood object of transitive verb
CL	classifier	ONE:MVMT	action done with one movement, suddenly
COM	comitative	P	postpositional or other oblique phrase
COMP	complementizer	PART	partitive
COMPLT	completive	PAST1	past tense from roughly one week to one month
COR	Set I coreferential clitic <i>jý-</i>	PAST2	past tense from roughly one month to one year
CORO	Set II coreferential clitic <i>-yu</i>	PAST3	distant past tense
COULD	'could' auxiliary (see also FRUST)	PERF	perfect
CU	modal clitic <i>-cu</i>	PL	plural
DAT	dative	PLEXCL	plural exclusive
DAY	phrasal clitic <i>-day</i>	PLINC	plural inclusive
DEMO	demonstrative root	PM	pragmatically marked component
DEPART	action done upon, or in preparation for, departure	POT	potential/optative
DIM	diminutive	PROX1	proximate 1 tense (earlier today or future)
DISTRIB	distributive	PROX2	proximate 2 tense (yesterday or future)
DL	dual	QUEST	yes/no question particle
DLEXCL	dual exclusive	REP	repetitive
ED	encoding device	S	subject; or single argument of intransitive verb
EMPH	emphatic clitic <i>-tée</i>	SG	singular
EP	phenomena to be encoded	SJL	San José de Loretoyacu dialect
EXCL	exclusive	TIY	clitic <i>-tiy</i>
FRUST	'frustrative' auxiliary (see also COULD)	TRNS	transitivizer (usually valence-increasing)
HABIT	habitual	V	verb; Vainilla dialect
IMPF	imperfective		
INAN	inanimate		
INCL	inclusive		
INF	infinitive/participial		
INST	instrument		
ITER	iterative		
ITER:MVMT	iterative movement		
IRR	irrealis auxiliary		
JIITA	second position clitic <i>jííta</i> or <i>jíí</i>		

xiv Abbreviations

VIN	verb initial norm (from	1	first person
	Keenan 1977 and 1979a)	2	second person
VRBLZR	verbalizer	3	third person

Chapter One

Introduction

Located in northeastern Peru, Yagua comes from an area of the world which has to date figured little in formulations of linguistic universals and theory construction. This study provides a typologically oriented description of aspects of the grammar of Yagua, as such pertain to constituent order type, constituent order co-occurrences, and discourse and pragmatic factors accounting for alternative orders.

Most right-thinking typologists would - I claim - consider Yagua to be a verb initial language at the level of major constituents of the clause. In this work we will first review and critique criteria for determining "basic word order" within the general typological tradition initiated by Greenberg (1963). Throughout the study evidence will be presented for and against the verb initial status of Yagua main clause structure. Given my claim that Yagua is a verb initial language, features which Edward Keenan has suggested commonly correlate with a consistent verb initial language type will first be summarized (Chapter 2). These syntactic and morphological features will then be discussed for Yagua (Chapters 3-7).

Main and subordinate clause morphosyntax is primarily addressed in Chapter 3. Relative to the noun phrase, Keenan suggests that in verb initial languages, there is commonly little, if any, agreement between modifiers and their head nouns. Noun phrase structure and agreement are explored in Chapters 4 and 5. In Yagua, the first issue to settle is what is a "modifier". The class of syntactic adjectives is extremely small - limited to perhaps one or two items. This hardly means that modifiers are nonexistent, however. Head noun versus modifying noun are distinguished on a discourse-pragmatic basis. In the process of exploring noun phrase structure and verb agreement, we will see that Yagua noun classification morphology provides a nice test case for S. Anderson's (1982) claim that a theoretically clear distinction between inflectional and derivational morphology can be maintained. The conclusion advanced here is that in some contexts the Yagua noun classification morphology must be accounted for synchronically by inflectional processes, but in other contexts it must be accounted for by derivational processes. A prototype view of inflection and derivation is explored and argued for. Chapter 6 explores relevant aspects of verb phrase structure.

Criteria commonly advanced for determining basic constituent order overlook the fact that in many, if not most languages of the world, transitive clauses rarely contain two overt noun phrase arguments, and

2 1. Introduction

then only under conditions which are marked relative to discourse-pragmatics. Chapter 7 discusses pragmatic factors motivating variations in order, and concludes that despite the scarcity of transitive clauses with two overt noun phrase arguments, the basic order within a classic typological tradition must be taken as verb-subject-object. The Yagua data suggest that Hawkins' (1983) proposed word order universals cannot be taken as exceptionless. At the present point in time Yagua is a highly "inconsistent" verb initial language. This may partly be a consequence of a historically prior OV order. Drawing partly on the work of Nichols (1986), some possible motivations are suggested for the particular directions of historical change which have resulted in the present conjunction of properties (Chapter 8).

1.1. Genetic and typological affiliations

Yagua is the only extant member of the Peba-Yaguan family. Formerly this family consisted of at least Peba, Yagua, and Yameo (Rivet 1911; Loukotka 1968). Rivet provides the only readily available Peba data, taken from colonial sources and largely limited to lexical items. Peba was spoken north of the town of Pebas on the Amazon river, north of the current Yagua area. Espinosa (1955) provides some information on Yameo, also largely limited to lexical items. Yameo was spoken in the region west of the jungle city of Iquitos, near the curve where the Amazon river changes course from north to east. Based on mass vocabulary comparison among numerous Amazonian languages, Rivet suggests that Peba-Yaguan is part of the Macro-Carib grouping. This is not well substantiated, however.

Greenberg (1960, 1987) also claims that Peba-Yaguan is a major branch of Macro-Carib, along with Huitotoan and Carib. Macro-Carib is purportedly a member of the Ge-Pano-Carib phylum. Scant evidence is presented for either of these claims. Loukotka (1968), Voegelin and Voegelin (1977), and Key (1979) follow Greenberg. For now I take an agnostic position on the larger genetic affiliation of Peba-Yaguan (though see Doris Payne 1984 and 1985c for one hypothesis).

Figure 1 shows the geographic location of Yagua relative to other language families in the Peruvian Amazon area.



Fig 1: Geographic Location of Yagua

No systematic study of shared typological traits among languages of the western Amazon has yet been undertaken.¹ Nevertheless, there are clear parallels between Yagua, Bora-Huitotoan, Zaparoan, Tucanoan, Cahuapana and some Maipurean Arawakan noun classification systems (Doris Payne 1987b). There are some striking similarities in verbal morphology and phonological processes with the Zaparoan languages, and more limited similarities in terms of noun classification systems (Doris Payne 1984, 1987b). Constituent order type (VSO, postpositional) parallels that of certain Maipurean Arawakan languages and of Taushiro (genetic affiliation uncertain) in the western Amazon, and of Guajajara (Tupí-Guaraní) in Brazil. A widespread South American is the infrequent use of noun phrases (Doris Payne 1986a). Another widespread feature may be a small or non-existent syntactic class of adjectives. Nominal modifiers are usually other nouns, but in natural discourse use of modifying words is infrequent. Use of nouns as modifiers is found in at least Yagua, Arawakan, Carib (e.g. Hixkaryana), Chayahuita (Cahuapana), and Quechua. There are similarities in the discourse environments motivating use of object clitic forms to refer to subjects in Yagua and at least some PreAndine Maipurean Arawakan languages (T. Payne 1985).² General organization of the verbal morphology is probably similar to, though not as complex as, that of the PreAndine Arawakan and Panoan languages, and of Cayuvava (genetic affiliation uncertain).

1.2. Demography and ethnography

The Yagua currently live in northeastern Peru in an area which P. Powlison (1969:3) describes as a rectangle approximately 200 miles wide and 350 miles long, extending between the second and fifth parallels and between the seventieth and seventy-fifth meridians. Chaumeil (1981) estimates that currently there are some 3000 Yaguas. Of these, Tom Payne (personal communication) estimates that roughly 75% of the women and 25% of the men are monolingual in Yagua, with the rest being bilingual in Spanish to varying degrees. Determination of the precise number of ethnic Yaguas is difficult due to ongoing assimilation into the mestizo culture and to long-standing social downgrading of the indigenous groups. If they can pass for mestizos, many ethnic Yaguas do not claim to be Yaguas.

Fejos (1943) is the first authoritative ethnographic study of the Yaguas, based on nine months of experience with them. (Tessmann 1930 gives some information based on second-hand reports; much of his information is incorrect.) Paul and Esther Powlison began living in the Yagua area in (1953), and have spent time with them intermittently until the present. P. Powlison (1969), a detailed study of Yagua folklore, contains the most accurate ethnographic description, including information on Yagua ceremonies and belief system. Even though a number of Yagua concentrations are currently located near the Amazon and other larger rivers of the region, theirs is traditionally a forest culture as opposed to a river culture. A large proportion of their daily food supply comes from cultivated chacras (swidden gardens), and now from fish; nevertheless, the Yaguas still consider the more arduous hunting task important. Chaumeil (1981) discusses pressures which have lead to this distinction between the preferred traditional hunting culture system and the fishing/horticultural system from which most of their actual food supply derives. (Reichel-Dolmatoff 1971 provides a fascinating discussion of similar ethnographic and culture distinctions among the Desana, a Tucanoan group to north.)

Chaumeil (1981) is a detailed tracing of Yagua migration patterns since the time of the early Jesuit missionaries in the 1700s until the present. Seiler-Baldinger (1976) gives additional information on some migrations near the Peruvian-Colombian border. The dialect situation has never been critically studied. However, the extensive migration within the last 80 years suggests that dialect differences cannot be adequately keyed to present-day geographical locations. Informal observations by Tom Payne and myself are that most differences pertain to phonetics and phonology, but there is also some morphological and minor syntactic

variation. Such variations will be noted where we are aware of them. Examples in this study come from three areas: San José de Loretoyacu (SJL) near the Peruvian-Colombian border, Cahocuma (CAH) north of Villacorta on the Amazon River, and Vainilla (V) near the confluence of the Napo and Amazon Rivers.

1.3. Previous linguistic work on Pebá-Yaguan

Chaumeil (1976) and Wise (1986) constitute a nearly exhaustive bibliography of published and microfiched material on Yagua and Pebá-Yaguan as of the mid 1970s (Wise covers linguistic work under the auspices of the Summer Institute of Linguistics up until about 1984). Chaumeil's bibliography includes colonial work of the early Jesuit missionaries dating from the 1700s. The Jesuit materials provide information on early western contact with numerous indigenous groups in the Amazon area. They also contain some short religious texts translated into Pebá and Yameo and a number of linguistic observations (usually from the perspective of Indo-European language structure). Following the Jesuits (1700s to early 1900s), the Benedictine (1800s to 1950s), Franciscan (early 1900s - 1920s), and Augustinian (early 1900s to mid 1950s) missionaries also left records of their contact with the Yaguas, Yameos, and Pebanos.

From a linguistic standpoint, perhaps the first important work is that of Rivet (1911). At least some, if not all, of Rivet's material is taken from colonial sources. It is largely limited to lists of lexical items and comparison of pronominal forms between Pebá, Yagua, and Yameo. Rivet felt that Pebá and Yagua were more closely related, as opposed to Yameo.

The third section of Espinosa (1955) gives more detailed linguistic information on Yameo. Espinosa's information was taken both from colonial sources and field work he did in the 1950s. At that time there were approximately 50 older speakers. Espinosa gives some information on verbal prefixes, adpositional phrases, and simple clause structure.

The first significant linguistic work on Yagua was done by Esther Powlison and Paul Powlison. Published linguistic work by the Powlisons consists of articles dealing with phonology (E. Powlison 1971, P. Powlison 1962), the number system (Powlison and Powlison 1958), and paragraph structure in a folktale (P. Powlison 1965). P. Powlison (1961) is a microfiched tentative grammar sketch which contains many useful observations about the meaning and distribution of various morphemes. Wise

(1986) lists additional unpublished microfiched materials by the Powlisons.

1.4. Data for the current study

The findings in this work are based on a corpus of well over 2,500 clauses of oral text, five short written texts, and field work carried out by Tom Payne and myself between February 1981 and April 1983. In addition, Paul Powlison made available his extensive text collection consisting of some 36 oral folkloric, personal narrative, and procedural texts (Powlison and Powlison 1977). A comprehensive morpheme concordance of these texts and preliminary dictionary materials consisting of some 3,000 entries have also been consulted.

A number of frequency counts and other observations made in Chapters 4 and 7 are based on exhaustive examination of 11 narrative texts, both oral and written. These are presented in Table 1.1. "Number of clauses" indicates the number of full clauses included in various counts from each text. The oral texts are divided into three groups: historical narrative, folkloric narrative, and personal narrative. The written texts are all personal narratives. Differences in subgenres do not appear to affect in any way the claims of this study.

Several comments are in order about the texts listed in Table 1.1. First, there is no well-established written tradition in Yagua. The Clausura text was actually spoken first in Spanish onto an audio cassette tape, after which it was translated via writing into Yagua by a more fluent speaker of Yagua. I thus consider it a basically written form, rather than an oral form. Quantification of different phenomena across the oral and written personal narrative groups shows no significant differences in the features compared. For instance, use of noun phrases across the two groups is statistically the same.³ Second, the Hunter's Text is not technically a personal narrative since it is not first person. However, it does not fit into either the historical nor the folkloric narrative type. Statistical comparison of different features with the Lechi Caspi text (which includes a fair amount of third person narration) shows no significant differences.

Table 1.1. Texts used for quantification of constituent orders and conditions for alternative orders

Genre	Text	Approximate number of clauses
Oral	Historical	Three Warriors
		David
		David Appendix
		<i>Total</i>
	Folkloric	First Squirrel
		Kneebite Twins
		Musmuqui
		<i>Total</i>
	Personal	Lagarto
		Hunter's Text
		Lechi Caspi
		<i>Total</i>
Written	Personal	Pq̄achi
		Clausura
		<i>Total</i>
	<i>Total clauses</i>	

The folkloric narrative group contains texts which are well known in the culture and which describe folk heroes. The texts partially explain how the world as known by the Yaguas came into being, and/or contain supernatural experiences. As a group these texts have fewer noun phrases and a higher incidence of intransitive clauses than the other genres.

The historical narratives do not, as a rule, provide a *raison d'être* for the world as it is or have the supernatural features characteristic of the folkloric narratives. They are old, probably widely-known stories, relating incidents that must have occurred 100 years ago or more in Yagua history. As a group, they have the highest percentage of noun phrases and the highest percentage of transitive clauses. The latter characteristic may be partially an artifact of the number of fighting events reported in these texts.

The main language consultants for this research were:

1. Pedro Díaz Cahuachi, age 18,⁴ from Urco Miraño, Peru. Pedro is the third son of Manungo Díaz, the traditional chief of the monolingual sector of the Urco Miraño community. The monolingual sector of this community migrated in the 1970s from Cahocuma, downriver on the Amazon from the town of Pebas. Consequently, Pedro's dialect is described as that of Cahocuma (CAH). He has had approximately three years of schooling in the local bilingual school, sponsored by the Peruvian Ministry of Education. Although he would consider himself bilingual, he is more at home in Yagua. He is married to an ethnic Yagua who professes to know almost no Yagua, and thus Spanish is possibly spoken in the home.⁵ His mother, sisters, and most of his extended family are monolingual in Yagua. His father is nearly so. When Pedro began to work with us, he had minimal literacy skills in Spanish and almost no experience reading or writing Yagua. Pedro gave us our first in-depth introduction to Yagua language and culture, and invited us to share in the building of his first house.

2. Hilario Peña Cahuachi, approximately 30-35 years of age, from Vainilla (V). Hilario has had considerable experience working with Paul Powlison on translation of the New Testament into Yagua, he is quite fluent in both Yagua and Spanish, and he has adequate literacy skills in both Yagua and Spanish. Hilario served as the language consultant for my most extensive research on the verbal morphology.

3. Mamerto Macahuachi, approximately 30-35 years of age, also of the Vainilla (V) dialect. Mamerto is a true bilingual and is more comfortable in the city and mestizo culture than our other language consultants. Perhaps because of his unusual degree of self-confidence and skills in both cultures, he was able to give us the first genuinely written texts in Yagua that we have been able to obtain.

4. Alcides Lozano Salazar, approximately 18 years of age, from San José de Loretoyacu (SJL). Alcides had six years of schooling in a local Spanish-speaking school when we first met him. (He has since received training as a bilingual teacher and has taught alongside a monolingual Spanish-speaking teacher.) Our linguistic work with Alcides was limited in duration, but helpful in discovering certain dialect distinctions. Alcides served as a language consultant on questions of the noun classification system.

In addition to the people specifically named, we interacted with a number of monolingual speakers of the Cahocuma dialect in Urco Miraño.

Chapter Two

Constituent Order and Order Correlations

The study of constituent order correlations began some 100 years ago, though it has been most intensely researched and theorized about during the last 30. In this chapter we will first review and critique major threads in the typological approach to this question, and then summarize some unpublished statements by Edward Keenan as to what is generally true of verb initial languages.

2.1. Observations of constituent order co-occurrences

According to Greenberg (1963:fn 4), the earliest reported observations of basic constituent order correlations come from the nineteenth century:

... the relation between genitive position and prepositions vs. postpositions and the hypothesis that some languages favor the order modifier-modified and others the opposite order is already a familiar notion in R. Lepsius' introduction to his *Nubische Grammatik* (Berlin, 1880).

Schmidt (1926) gives a more studied treatment of selected orders based on a world sample. Greenberg (1963:83) summarizes Schmidt's basic conclusions:

Prepositions go with nominative-genitive order and postpositions with the reverse order. The nominative-genitive order tends to appear with verb before nominal object and genitive-nominative with object-verb... Further, nominative-genitive is associated with noun-adjective and genitive-nominative with adjective-noun.

Greenberg 1963 is, of course, a landmark work on observations of constituent order co-occurrences. Based on a sample of 30 languages, he draws a number of statistical and absolute implicational universals. These range over a wide variety of syntactic and morphological features. Appendix II to Greenberg (1963), based on a more extensive language sample, lists 24 possible combinations of subject-object-verb, adpositional

phrase, noun + genitive, and adjective + noun orders. Perusal of this Appendix shows that some combinations are heavily attested in the sample, while others are not found at all. He nevertheless cautions that the proposed universals are to be taken as tentative, pending a more complete sample. This is an important caution. For instance, Universal 3 states (88): "Languages with dominant VSO order are always prepositional". However, Hawkins (1979, 1983) attributes to Keenan the more recent observation that this universal does admit of some exceptions. Keenan's statement is at least partly based on South American Arawakan languages, as Keenan (1978:292) notes that Baure (Bolivian Arawakan) and other related languages are verb initial plus postpositional.

Hawkins 1979, 1980, 1982a, and 1983 are extensions of Greenberg's work, based on a sample of some 350 languages. This extended sample shows generally similar attestation of co-occurrence types as does Greenberg's Appendix II. Nevertheless, Hawkins apparently did not pick up on the VSO/V-initial plus postpositional combination as an Arawakan pattern, as he cites Pima-Papago (Uto-Aztecan) as the only attested example of a VSO-postpositional language (but see Doris Payne 1987a for evidence against classifying Papago as VSO). In addition to studying distribution and co-occurrence of adposition, subject-object-verb, noun-descriptive modifier (adjective), and genitive-noun orders, Hawkins (1983) also explores co-occurrence orders of other constituents within the noun phrase. Some specific issues that the Yagua data raise for Hawkins' proposals are addressed in Section 2.3 and Chapter 8. Now, however, we turn to correlations specific to verb initial languages.

2.2. The verb initial norm (VIN)

As far as I know, there is no published statement of features typically found in verb initial languages. Keenan 1978 on the syntax of subject final languages is perhaps the nearest approximation to such a statement. In this section, I reproduce a number of observations extracted from Keenan's (1977) "Summary of word order typology", and from his 1979a manuscript on "Word order typologies: the verb initial typology". I have recast the observations in complete sentences and made other changes of an editorial nature. Throughout the remainder of this work I refer to these observations as the "Verb Initial Norm" (VIN).

1. General. Verb initial languages are largely, though not entirely, the mirror image of verb final languages.

2. Morphology

2.1. Verb initial languages evidence significant prefixing, though normally there is some suffixing as well. There is a possibility of ambi-fixing (discontinuous affixes), and a somewhat greater than chance tendency for discontinuous demonstratives.

2.2. Verb initial languages may be agglutinative and polysynthetic.

3. Basic word order

3.1. Verb initial languages are comprised of the following types:

- [1]. Verb initial plus free order of full NPs
(Tagalog)
 - [2]. V-DO-S-Obl (Fijian, Toba Batak)
 - [3]. V-DO-Obl-S (Malagasy, Tzeltal)
 - [4]. V-S-DO-Obl (Celtic, Eastern Nilotic,
Polynesian, Jacaltec)
- Type [4] is by far the most common.

3.2. Freedom. Fronting of subject NPs to the left of the verb is always a possibility, though often it is morphologically marked in some way (not necessarily on the NP). The order after the verb is frequently rigid, though sometimes quite free as in Tagalog and, to a lesser extent, in Chinook.

4. Sentence level syntax

4.1. Topicalization. Topicalization may be done by fronting, though there is a tendency in Nilotic to move old information to the end of the clause.

4.2. Focussing. Focussing of information as in a cleft or information question is done by fronting. Often this may be accompanied by particles separating the subject from the rest of the clause. The result is always pragmatically marked, i.e. emphatic, contrastive, focussed, etc.

4.3. Comparisons. The comparative form precedes the standard. The comparative marker is commonly a verbal form, or else an adposition.

Thus, *John is taller than Bill* may be expressed as *Tall John from-Bill*, or as *Tall John exceed Bill*.

4.4. Questions

4.4.1. In yes-no questions the question particle, if any, occurs sentence initially.

4.4.2. In NP questions, a questioned NP is always frontable and this is the normal pattern. It is possible, but less normal, to leave the questioned NP in the position questioned. A few cases of rightward movement of question words are attested, but there is no attested tendency for the question word to attract to the normal DO position (as is the case for verb final languages).

4.5. Subordinate clauses and sentence complements

4.5.1. It is very common for many types of subordinate clauses to be finite.

4.5.2. Subordinating markers such as complementizers, nominalizers, and subordinate conjunctions precede their clauses.

4.5.3. Sentences which are subordinate to verbs, adjectives, or nouns invariably follow the element to which they are subordinate.

4.5.4. Adverbial subordinate clauses usually follow their main clauses. For example *Will leave John because is tired Mary* occurs for *John will leave because Mary is tired*. However, frontability of conditionals is likely universal (cf. Greenberg 1963).

4.6. Coordinate sentences are commonly expressed as [S and S]. [S, S and] is not attested. Perhaps the existence of overt coordinate conjunctions at the S level, especially *or*, is less well attested than in verb medial languages.

4.7. Speech act indicators (e.g. question particles, etc.) are normally sentence initial, though other positions are possible.

5. The noun phrase

5.1. Case marking

5.1.1. All major NPs may be case marked (Tongan, Nandi), but it is very common for most major NPs to carry little or no nominal case marking. Where affixal case marking occurs, it is more likely to be prefixal than in verb final languages, but suffixing is still fairly common.

5.1.2. Where case marking exists it is normally done by prepositions (though some Amerindian languages are exceptions here, such as Machiguenga and Quileute, which have postpositions).

5.1.3. Verbal case marking is attested to a very significant degree. That is, verbs carry affixes indicating that an instrumental, goal locative, benefactee, etc. is present, and the corresponding full NPs carry no adpositions or distinctive case marking.

5.1.4. As with verb final languages, but in distinction to verb medial languages, case marking (and verb agreement) may follow an ergative pattern.

5.2. Adjectives

5.2.1. The demonstrative, numeral, and qualifying adjective follow the common noun in that order or its mirror image (Adj + Num + Dem).

5.2.2. There is probably less agreement with common nouns than in verb final languages, especially case agreement.

5.2.3. Adverbs follow adjectives (but this needs further checking).

5.3. Articles

5.3.1. The presence of definite articles distinct from demonstratives is much more common than in verb final languages.

5.3.2. The existence of several articles (definite, indefinite, specific, plural, proper noun) is much more common than in verb final languages (e.g. Maori, Fijian).

5.4. Possessors: With great regularity Possessor NPs follow the head NP, as in *father of John* rather than *John's father*.

5.5. Relative clauses

5.5.1. The dominant order is always postnominal.

5.5.2. Occurrence of personal pronouns in positions relativized is fairly common, though relativization by deletion is still the most common strategy.

5.5.3. In distinction to verb final languages, co-relatives are not attested.

5.5.4. Like verb final languages, but in distinction to verb medial languages, relative pronouns which code the case of the position relativized are rare. It is less rare than in verb final languages, however (e.g. Tamazight, Berber).

5.5.5. Relative pronouns which agree with the head noun in noun class and sometimes even case are attested (e.g. Classical Arabic, Nandi).

5.5.6. In distinction to verb final languages, internally headed relatives are not attested, though the phenomenon is not well studied.

6. The verb phrase

6.1. Tense/aspect, passive, inchoatives, causatives, negation, modals, desideratives and volitionals may appear marked on the verb. There is significantly more prefixing in verb initial languages than in verb final ones, and very possibly more ambifixing and infixing. There is, to Keenan's knowledge, always some suffixing, however.

6.2. If expressed by morphemically independent forms, modals, auxiliaries (if such exist), negative particles or words, desideratives and volitionals always precede the main verb, and may themselves have independent verbal morphology. (This may also be true for tense/aspect, passive, inchoatives, and causatives.) The strength of the order correlation here is better than its converse for verb final languages.

6.3. Manner adverbs follow the verb if they are a distinct category (which often they are not).

6.4. Sentential objects always follow the subject and are very commonly finite as opposed to the more usual non-finite/nominalized treatment they receive in verb final languages.

6.5. Sentential objects are never embedded. They normally follow the main sentence but may precede, especially in direct quote contexts.

6.6. Verbal forms subordinate to the "main" verb (e.g. complements of verbs like *want*, *try*, etc.) always follow the main verb, and are commonly finite.

6.7. Causativized verbs follow the causativizing verb.

6.8. "Backward" equi-deletion may occur. That is, *want John go* or *want-go John* may occur for 'John wants to go'. This is never a possibility in verb final languages.

6.9. There is possibly less rich means for nominalizing and definitizing verb phrases than in verb final languages. On the other hand, in many but not all verb initial languages the verbal complex seems historically to be a nominal construction, at least in part (Middle Egyptian, Welsh, Malagasy, Philippine languages, Mayan).

6.10. Verb initial languages always have a passive voice and it is almost always marked in the verbal morphology (rather than by a serial verb construction as in Chinese, for example). It may be marked by a verb plus nominalization as in *John receive hitting from Bill* (Tzeltal, Mayan).

6.11. With possibly greater than chance frequency, the verb in verb initial languages either agrees with no NPs, or with two NPs (both subject and direct object, or sometimes subject and indirect object).

6.12. Verb initial languages normally have no overt copula.

In the following chapters, the Yagua data will be compared with this verb initial norm. Yagua proves to be very mixed typologically, though it evidences more than half of the characteristic verb initial traits.

2.3. Selected theoretical approaches accounting for word order correspondences

The preceding section reviewed selected order and morphosyntactic correlations noted for verb initial languages. Why certain correlations should tend to occur must derive from psycholinguistic constraints on, and pressures towards, morphosyntactic change, coupled with the grammaticization of certain discourse-pragmatic frequency patterns.

Greenberg did not propose a unified theory accounting for his observed universals. He did, however, reflect in important ways on his observations. The operator (modifier) - operand (modified) distinction is commonly attributed to Lehmann and Vennemann (cf. Lehmann 1973; Vennemann 1974; Vennemann and Harlow 1977); but Greenberg (1963) and Lepsius before him noted that in most languages there is a tendency to put either the modified element before the modifier, or vice versa. Greenberg also noted the greater cross-linguistic ambivalence of adjective - noun order, which he attributed to analogies with other constructions. Based on a sample of 506 languages Dryer (1986), in fact, argues that there is no evidence of any relationship between order of verb and object, and order of noun and adjective. Similarly, the seeds of Hawkins' Cross-Category Harmony principle (cf. Hawkins 1982a, 1983) are found in Greenberg's discussion of harmonic and disharmonic relations among distinct rules of order, presumably associated with psychological generalization. Hawkins throws out SVO as a distinctive type, noting that nothing specifically correlates with SVO order. But Greenberg had earlier stated (79): "One may further conjecture that if there are exceptions they will be in type II [SVO], which, having both SV and VO which are disharmonic, can provide an anchor in either case for deviant genitive order".

Lehmann (1973) and Vennemann (1974, 1975, 1981) have theorized about the principles underlying Greenberg's observations. Their proposals are based on the modifier-modified distinction, which is extended to provide diachronic explanations of constituent order change. Lehmann (1973) makes a broad distinction between OV and VO languages, and is principally concerned with an ordering principle governing placement of modifiers relative to their heads in "consistent" languages (48): "modifiers are placed on the opposite side of a basic syntactic element from its primary concomitant". Thus, in OV languages, relative clauses, adjectival, and genitival expressions precede their head nouns, since the primary concomitant of the (object) noun is the following verb. In VO languages, relative clauses, adjectival, and genitival expressions follow their heads for the same reason. For Lehmann, then, there is no distinc-

tion between SVO, VSO, and VOS types, as all are VO. As modifiers become affixal through phonological reduction, the ordering principle supposedly leads to suffixal agglutinative morphology in consistent OV languages, but to prefixing morphology in consistent VO languages. Lehmann suggests that there is a tendency for VO languages to be more isolating or inflectional due to the disruptive influence of the subject following the verb (however, note that this says nothing about why an SVO or VOS language might be isolating or inflectional - both of which are VO). Languages which are not consistently OV or VO are assumed to be in the process of historical change. However, no cogent reasons are given as to how or why inconsistency might be introduced to begin with, or for the huge number of inconsistent languages which have been in their "unstable" state for centuries.

In addition to an overly simplistic division between OV versus VO languages and problems with historical change, a potential difficulty with Lehmann's principle is the notion "primary concomitant of a verb". He assumes a theory of universal grammar containing phrase structure rules in which the sentence S consists of two initial components. Among the early phrase structure rules is the rule $S \rightarrow QP$, where Q stands for Qualifier (sentence constituents which modify the entire proposition), and P stands for Proposition. I infer that in Lehmann's schema, some subsequent rule exists such as $P \rightarrow VN(=OBJ)$, where V and N(=OBJ) are unordered relative to each other (cf. 1973:49). Lehmann explicitly rejects inclusion of initial phrase structure rules such as (a) $S \rightarrow NPVP$ and (b) $VP \rightarrow VNP$, where rule (a) introduces a subject phrase as a primary element along with the verb phrase (51). His reasons for excluding rules such as (a) from universal grammar are that (1) subjects are (often) not mandatory or "primary elements in sentences", as in Japanese and Hebrew; (2) their inclusion as primary elements has resulted in "trouble for typologists as well as for linguistic theorists in general" as they have tried to "classify SVO and VSO languages as major types in the same way as VO and OV languages"; and (3) there is the problem of languages where the identification of a single nominal as "subject" is problematic. Thus, Lehmann rejects consideration of subject nominals as "primary concomitants" of the verb phrase.

Considering these arguments against inclusion of the subject as a primary concomitant of the verb phrase, we might well ask why the object constituent should be considered a "primary concomitant" of the verb in terms of universal grammar in general, and of Yagua in particular. Although there is little or no problem in identifying subject versus object nominals in Yagua (argument 3 above), it is certainly not the case that identification of the syntactic role "object" is non-problematic world-

wide (cf. Schachter 1984 on Toba Batak for one such language). Further, in terms of frequency, objects are not "mandatory" in Yagua clauses in context (cf. Chapter 7), and this is true in a number of other languages as well (cf. Derbyshire 1979, 1986b; Scancarelli 1985; Du Bois 1987; Doris Payne 1986a; Wise 1986). In Yagua the only mandatory elements are the verb or predicate nominal, plus clitic reference to the subject and/or object argument (the subject clitic is syntactically obligatory only if there is no full NP, and optional otherwise). After the verb, the next most "mandatory" element in Yagua would be a postpositional phrase or an adverbial element. As I will suggest in Chapter 7, there are certain difficulties and indeterminacies in trying to substantiate that V(S)O is any more "basic" a clause type than simply V + clitic; in terms of discourse, V + clitic may in one sense be more neutral and communicatively basic. Further, the existence of VSO (and possibly OSV) languages generally, where the verb and object are not necessarily contiguous, raises other questions as to why the object should universally be considered the "primary concomitant" of the verb in a structural sense.¹

Vennemann (1974) accepts Lehmann's distinction between OV and VO types, and proposes the Natural Serialization Principle. This claims that "consistent" languages will serialize all operators (modifiers) to one side of their operands (heads). The NSP is a bivalued and implicational statement of the form: if P, then Q ($P \rightarrow Q$). It is reversible: $P \rightarrow Q$, and $Q \rightarrow P$. For example, if OV, then postpositional; and if postpositional, then OV (where adposition and verb are operands, and NP and O are operators). As there are numerous languages which stand as exceptions to such strong claims, the NSP is presented as a statistical principle, defining preferred consistent types. Relative to diachrony, inconsistent languages are supposedly moving from one consistent type to another; to a great extent, verb position is taken as the trigger to which other operand orders will conform over time. Operand status is determined by two factors: (1) If syntactic category constancy is maintained between a constituent X of a phrase, and the phrase XP itself, then X is the operand. (2) A logico-semantic criterion stipulates that operators are those elements which specify (i.e. are functions on) operands.

Hawkins (1980, 1983) provides a good critique of the inadequacies and logical inconsistencies in Vennemann's proposals. First, Vennemann's definition of operand versus operator is based on a logical argument-function distinction, but Keenan (1979b) argues that Vennemann's operator-operand constructions do not correspond to standard logical function-argument distinctions. Hawkins concludes that the operand-operator (modified-modifier) distinction IS the significant level of generality for serialization principles (including the NSP); attempts to trace them back