The Amazonian Languages

edited by
R. M. W. Dixon and
Alexandra Y. Aikhenvald
THE AMAZONIAN LANGUAGES

The Amazon Basin is arguably both the least-known and the most complex linguistic region in the world today. It is the home of some 300 languages belonging to around 20 language families, plus more than a dozen genetic isolates, and many of these languages (often incompletely documented and mostly endangered) show properties that constitute exceptions to received ideas about linguistic universals. This book is the first to provide an overview in a single volume of this rich and exciting linguistic area. The editors and contributors have sought to make their descriptions as clear and accessible as possible, in order to provide a basis for further research on the structural characteristics of Amazonian languages and their genetic and areal relationships, as well as a point of entry to important cross-linguistic data for the wider constituency of theoretical linguists.

R. M. W. DIXON and ALEXANDRA Y. AIKHENVALD are Director and Associate Director of the Research Centre for Linguistic Typology at the Australian National University (the Research Centre will relocate to La Trobe University in Melbourne from January 2000). Professor Dixon’s book publications include grammatical studies of the Australian languages Dyirbal and Yidiny, of Fijian and of English, as well as The Languages of Australia (1980) and Ergativity (1994). One of his current projects is a grammar of Jarawara (Brazil). Professor Aikhenvald has published 6 books and nearly 100 papers in Russian, English, Portuguese and Yiddish, covering a range of subjects including Berber, Hebrew, Indo-European and Native South American languages. Her monograph Classifiers: A Typology of Noun Categorization Devices will be published in 1999. She is currently completing a full-length grammar of Tariana.
This series offers general accounts of the major language families of the world, with volumes organized either on a purely genetic basis or on a geographical basis, whichever yields the most convenient and intelligible grouping in each case.

Each volume compares and contrasts the typological features of the languages it deals with. It also treats the relevant genetic relationships, historical development and sociolinguistic issues arising from their role and use in the world today. The books are intended for linguists from undergraduate level upwards, but no special knowledge of the languages under consideration is assumed.

Volumes such as those on Australia and the Amazon Basin are also of wider relevance, as the future of the languages and their speakers raises important social and political issues.

Already published:
The languages of Australia R. M. W. Dixon
The languages of the Soviet Union Bernard Comrie
The Mesoamerican Indian languages Jorge A. Suárez
The Papuan languages of New Guinea William A. Foley
Chinese Jerry Norman
The languages of Japan Masayoshi Shibatani
Pidgins and creoles (Volume I: Theory and Structure; Volume II: Reference survey) John A. Holm
The Indo-Aryan languages Colin Masica
The Celtic languages edited by Donald MacAulay
The Romance languages Rebecca Posner
The Korean language Ho-Min Sohn
THE AMAZONIAN LANGUAGES

edited by
R. M. W. DIXON
and
ALEXANDRA Y. AIKHENVALD
Research Centre for Linguistic Typology
La Trobe University, Melbourne

CAMBRIDGE UNIVERSITY PRESS
CONTENTS

List of maps
List of contributors
Acknowledgements
List of abbreviations
Conventions followed

1 Introduction by R. M. W. Dixon and Alexandra Y. Aikhenvald

1 Cultural background 3
2 Linguistic diffusion 7
3 Genetic relationship 11
4 The punctuated equilibrium model 16
5 Organization of this book 19
Bibliography 20

2 Carib by Desmond C. Derbyshire

1 Introduction 23
2 Phonology 26
2.1 Segmental 26
2.2 Phonotactics and suprasegmentals 26
2.3 Morphophonology 28
3 Morphology, particles and pronouns 31
3.1 Inflectional morphology 31
3.1.1 Person-marking affixes on verbs, nouns, adverbials and postpositions 32
3.1.2 Tense, aspect, mode and number suffixes on verbs 37
3.1.3 Possession, tense and number suffixes on nouns 40
3.1.4 Inflectional suffixes on locative postpositions 42
3.2 Derivational morphology 43
3.2.1 Verb derivational affixes 44
3.2.2 Nominalizing affixes attached to verb stems to form nouns 45
CONTRIBUTORS

Alexandra Y. Aikhenvald
Research Centre for Linguistic Typology
La Trobe University
Bundoora, Vic, 3083, Australia
a.aikhenvald@latrobe.edu.au

Janet Barnes
ILV - Apartado Aéreo 120308
Bogotá, Colombia
Janet.Barnes@sil.org

Desmond C. Derbyshire
56 Furze Hill Road
Headley Down
Burdon, GU35 8HA, England
des_derbyshire@sil.org

R. M. W. Dixon
Research Centre for Linguistic Typology
La Trobe University
Bundoora, Vic, 3083, Australia

Raquel Guirardello
Rua Ferrucio Celani No. 152
Centro - Valinhos - S.P.
13270-000 Brazil

Cheryl Jensen
Summer Institute of Linguistics
Ag. Cabanagem CP 5040
66601-970 Belem, Pará, Brazil
Cheryl_Jensen@sil.org

Eugene E. Loos
7407 Cave Drive
Dallas, TX 75249, USA
eeloos@cyberramp.net

Ivan Lowe
Summer Institute of Linguistics
Horsleys Green, High Wycombe
Bucks., HP14 3XL, England
ivan-margaret_lowe@wycliffe.org

Silvana and Valteir Martins
Rua Guaiaranazes 357
Bairro Paraiso
16001-970 Araçatuba, S.P. Brazil

Aryon D. Rodrigues
SQN 410 K 102
70865-110 Brasília, DF, Brazil

List of contributors

Lucy Seki
Universidade Estadual de Campinas
Instituto de Estudos da Linguagem
Departamento de Linguística,
Cx. P. 6045
13081-970 Campinas, S.P. Brazil
lseki@iel.unicamp.br

Mary Ruth Wise
Summer Institute of Linguistics
7500 W. Camp Wisdom Rd,
Dallas, TX 75236-5699, USA
MaryRuth_Wise@sil.org
ACKNOWLEDGEMENTS

The editors would like to thank the linguistic scholars of South America (in universities, museums and missionary organizations) who have discussed with them the wonders of the South American languages, and freely shared their materials and ideas. And the speakers of Amazonian languages who have extended friendship; explained what to say to who in what circumstances and why; and helped unravel the intricate structures of their native tongues.

This volume could not have come into being without the critical assistance of two members of the Research Centre for Linguistic Typology. Jennifer Elliott, the Centre's Administrator, solved computational conundrums and restored phonetic symbols that had strayed. Suzanne Kite, our Research Assistant, spent hundreds of hours checking the chapters for consistency, style and sense; e-mailing contributors to clarify what they meant and the most appropriate way to express it; and integrating what began as a number of heterogeneous parts into a reasonably unified whole.

ABBREVIATIONS

The following abbreviations are used for the states of Brazil:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>State Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Acre</td>
</tr>
<tr>
<td>AM</td>
<td>Amazonas</td>
</tr>
<tr>
<td>AP</td>
<td>Amapá</td>
</tr>
<tr>
<td>BA</td>
<td>Bahia</td>
</tr>
<tr>
<td>ES</td>
<td>Espírito Santo</td>
</tr>
<tr>
<td>GO</td>
<td>Goiás</td>
</tr>
<tr>
<td>MA</td>
<td>Maranhão</td>
</tr>
<tr>
<td>MG</td>
<td>Minas Gerais</td>
</tr>
<tr>
<td>MS</td>
<td>Mato Grosso do Sul</td>
</tr>
<tr>
<td>PA</td>
<td>Pará</td>
</tr>
<tr>
<td>PR</td>
<td>Paraná</td>
</tr>
<tr>
<td>RJ</td>
<td>Rio de Janeiro</td>
</tr>
<tr>
<td>RO</td>
<td>Rondônia</td>
</tr>
<tr>
<td>RS</td>
<td>Rio Grande do Sul</td>
</tr>
<tr>
<td>SC</td>
<td>Santa Catarina</td>
</tr>
<tr>
<td>SE</td>
<td>Sergipe</td>
</tr>
<tr>
<td>SP</td>
<td>São Paulo</td>
</tr>
<tr>
<td>TO</td>
<td>Tocantins</td>
</tr>
</tbody>
</table>

The following abbreviations are used in the bibliographies:

IJAL  International Journal of American Linguistics
SIL   Summer Institute of Linguistics

Abbreviations used in interlinear glosses are:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>1+2</td>
<td>first person inclusive</td>
</tr>
<tr>
<td>1+3</td>
<td>first person exclusive</td>
</tr>
<tr>
<td>1R</td>
<td>relating to first person</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>A</td>
<td>transitive subject</td>
</tr>
<tr>
<td>ABS</td>
<td>absolutive</td>
</tr>
<tr>
<td>AC</td>
<td>action</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADJ</td>
<td>adjective</td>
</tr>
<tr>
<td>ADV</td>
<td>adverbal</td>
</tr>
<tr>
<td>ADVZ</td>
<td>adverbializer</td>
</tr>
<tr>
<td>AG</td>
<td>agentive</td>
</tr>
<tr>
<td>AG.NMLZR</td>
<td>agent nominalizer</td>
</tr>
<tr>
<td>ALIEN</td>
<td>alienable</td>
</tr>
<tr>
<td>ALL</td>
<td>allative</td>
</tr>
<tr>
<td>ANIM</td>
<td>animate</td>
</tr>
<tr>
<td>APPLIC</td>
<td>applicative</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associated</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>ATTRIB</td>
<td>attributive case</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary verb</td>
</tr>
<tr>
<td>BEN</td>
<td>benefactive</td>
</tr>
<tr>
<td>C</td>
<td>consonant</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
</tr>
<tr>
<td>CIR,NMLZR</td>
<td>circumstantial nominalizer</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CNT</td>
<td>contiguous marker</td>
</tr>
<tr>
<td>COLL</td>
<td>collective</td>
</tr>
<tr>
<td>COMIT</td>
<td>comitative</td>
</tr>
<tr>
<td>COMPL</td>
<td>completive aspect</td>
</tr>
<tr>
<td>COND</td>
<td>conditional</td>
</tr>
<tr>
<td>CONJ</td>
<td>conjunction</td>
</tr>
<tr>
<td>CONT</td>
<td>continuative aspect</td>
</tr>
<tr>
<td>COREF</td>
<td>coreferential</td>
</tr>
<tr>
<td>CURR</td>
<td>current</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative mood</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>DEIC</td>
<td>deictic</td>
</tr>
<tr>
<td>DENMLZR</td>
<td>denominalizer</td>
</tr>
<tr>
<td>DESID</td>
<td>desiderative</td>
</tr>
<tr>
<td>DET</td>
<td>determinate</td>
</tr>
<tr>
<td>DETRZR</td>
<td>detransitivizer</td>
</tr>
<tr>
<td>DIF,LOC</td>
<td>diffuse locative case</td>
</tr>
<tr>
<td>DIM</td>
<td>diminutive</td>
</tr>
<tr>
<td>DIST</td>
<td>distant</td>
</tr>
<tr>
<td>DL, dl</td>
<td>dual</td>
</tr>
<tr>
<td>DS</td>
<td>different A/S referent</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EVID</td>
<td>evidential</td>
</tr>
<tr>
<td>EXCL, excl</td>
<td>exclusive</td>
</tr>
<tr>
<td>EXISTS</td>
<td>existential</td>
</tr>
<tr>
<td>FEM, fem</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>IMM</td>
<td>immediate</td>
</tr>
<tr>
<td>IMPER</td>
<td>imperative</td>
</tr>
<tr>
<td>IMPERFV</td>
<td>imperfective</td>
</tr>
<tr>
<td>IMPERS</td>
<td>impersonal</td>
</tr>
<tr>
<td>INAL</td>
<td>inalienable</td>
</tr>
<tr>
<td>INAN</td>
<td>inanimate</td>
</tr>
<tr>
<td>INCL, incl</td>
<td>inclusive</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INDIC</td>
<td>indicative</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive/pseudo-infinitive</td>
</tr>
<tr>
<td>INST</td>
<td>instrument</td>
</tr>
<tr>
<td>INTERROG</td>
<td>interrogative</td>
</tr>
<tr>
<td>INTR</td>
<td>intransitive</td>
</tr>
<tr>
<td>IRR</td>
<td>irregular</td>
</tr>
<tr>
<td>LINK</td>
<td>linking or relational</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>MAL</td>
<td>malefactive</td>
</tr>
<tr>
<td>MEDIAL</td>
<td>medial</td>
</tr>
<tr>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>NCNT</td>
<td>non-contiguous determiner</td>
</tr>
<tr>
<td>NEG</td>
<td>negative/hegation</td>
</tr>
<tr>
<td>NEUT</td>
<td>neuter</td>
</tr>
<tr>
<td>n.fem</td>
<td>non-feminine</td>
</tr>
<tr>
<td>NMLZR</td>
<td>nominalizer</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NOM.CASE</td>
<td>nominal case</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>O</td>
<td>transitive (direct)</td>
</tr>
<tr>
<td>OBBL, OBL</td>
<td>oblique</td>
</tr>
<tr>
<td>OBL-TOP</td>
<td>oblique-topicalized</td>
</tr>
<tr>
<td>PART,LOC</td>
<td>partitive locative case</td>
</tr>
<tr>
<td>PASSV</td>
<td>passive</td>
</tr>
<tr>
<td>PAST.1</td>
<td>first past tense</td>
</tr>
<tr>
<td>PAST.2</td>
<td>second past tense</td>
</tr>
<tr>
<td>PERF</td>
<td>perfect</td>
</tr>
<tr>
<td>PERFV</td>
<td>perfective</td>
</tr>
<tr>
<td>PL, pl</td>
<td>plural</td>
</tr>
<tr>
<td>POSSD</td>
<td>possession/possessed</td>
</tr>
<tr>
<td>POSSR</td>
<td>possessor</td>
</tr>
<tr>
<td>POSSV</td>
<td>possessive</td>
</tr>
<tr>
<td>PRES</td>
<td>present</td>
</tr>
<tr>
<td>PRO</td>
<td>pronoun</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive aspect</td>
</tr>
<tr>
<td>PROX</td>
<td>proximate</td>
</tr>
<tr>
<td>PUNCT</td>
<td>punctual aspect</td>
</tr>
<tr>
<td>PUNCT.LOC</td>
<td>punctual locative case</td>
</tr>
<tr>
<td>REAL</td>
<td>realis</td>
</tr>
<tr>
<td>REC</td>
<td>recent</td>
</tr>
<tr>
<td>RECIP</td>
<td>reciprocal</td>
</tr>
<tr>
<td>REL</td>
<td>relative clause marker</td>
</tr>
<tr>
<td>RELZR</td>
<td>relativizer</td>
</tr>
<tr>
<td>REM</td>
<td>remote</td>
</tr>
<tr>
<td>REP</td>
<td>reportative</td>
</tr>
<tr>
<td>REPET</td>
<td>repetitive</td>
</tr>
<tr>
<td>RES</td>
<td>resultative</td>
</tr>
<tr>
<td>REV</td>
<td>reversative</td>
</tr>
<tr>
<td>SEQ</td>
<td>sequential</td>
</tr>
<tr>
<td>SER.V</td>
<td>serial verb</td>
</tr>
<tr>
<td>SF</td>
<td>stem formative</td>
</tr>
<tr>
<td>SG, sg</td>
<td>singular</td>
</tr>
<tr>
<td>SO</td>
<td>A/S referent same as O referent</td>
</tr>
<tr>
<td>SPEC</td>
<td>specificity marker</td>
</tr>
<tr>
<td>SS</td>
<td>same A/S referent</td>
</tr>
<tr>
<td>TAMN</td>
<td>tense, aspect, mood and number</td>
</tr>
<tr>
<td>TR</td>
<td>transitive</td>
</tr>
<tr>
<td>V</td>
<td>verb/vowel</td>
</tr>
<tr>
<td>VBLZR</td>
<td>verbalizer</td>
</tr>
</tbody>
</table>
CONVENTIONS FOLLOWED

I SPELLING

We have tried to avoid spelling conventions that are particular to one language, using instead those which are most generally accepted.

(a) Spanish 'hu' and 'j'. The sounds [w] and [h] are shown by 'hu' and 'j' respectively in Spanish orthography. We have generally preferred plain 'w' and 'h', e.g. Witoto rather than Huitoto, Cashinawa rather than Cashinhua, Guahibo rather than Guajuibo. (We have allowed just a few exceptions, e.g. Jivaro [Hivaro] is always spelt with 'J' and this has been retained. Cahuapana and Chayahuita are almost always spelt with 'hu' and this is retained here.)

The sound [k] is variously written as 'k' or 'c' or 'qu'. For names that include a [k] we have, as a rule, retained the spelling which is best known.

(b) English '-an'. English-speaking linguists often put '-an' on the end of the name of a language family, e.g. Arawakan, Cariban, Tupian. (Although this is not done consistently. Mercifully we have never seen Jéan or Makúan.) Spanish- and Portuguese-speaking linguists avoid names ending with an English suffix. We think it most useful to have one universal name for each family (rather than a certain name when writing in English and another name when writing in Spanish or Portuguese), and so avoid the '-an'. This does mean that sometimes a language family and one language within the family may be referred to by the same label, but context is almost always sufficient to avoid confusion.

There are additional reasons for following this convention. The term 'Arawak' is used by Spanish- and Portuguese-speaking linguists for a well-established language family. A number of English-speaking linguists use 'Arawakan' for something quite different, a chimerical high-level grouping (see §3 of chapter 1) which is said to include what South American linguists term 'Arawak' but is in this classification named 'Maipuran'. We follow the traditional (and current South American) practice of using 'Arawak' (not 'Maipuran') for the established language family, and avoid the term 'Arawakan'.

2 LANGUAGE AND DIALECT, FAMILY AND SUBGROUP

(a) The term 'language' can be used with two quite different meanings. One is the linguistic sense, when forms of speech that are mutually unintelligible are designated to be distinct languages. Two forms of speech that are mutually intelligible are then dialects of one language. The other is the political sense, when each separate nation or tribe likes to say that it has its own language. The two senses of 'language' often give the same result but sometimes they do not. One linguistic language may relate to several political languages, as when the Swedish and Norwegian peoples each say that they have a distinct language, whereas these are mutually intelligible and are thus, on linguistic grounds, dialects of a single language. Or one political language may relate to several linguistic languages as when the Chinese say that they all speak the Chinese language, but in a number of dialects. In fact the 'dialects' are not mutually intelligible and are, on linguistic criteria, separate languages.

In this book we employ the term 'language' in the linguistic sense. Note that earlier work on South America has used the term 'language' sometimes in the linguistic and sometimes in the political sense. For instance, there are seven Makú tribes, each of which has its own political language, but there are just four separate languages in the linguistic sense. Yanomami has sometimes been described as involving four closely related languages, but it seems most appropriate to characterize Yanomami as a dialect continuum (effectively, as a single language). Within the Arawá family, Jamamadi, Jarawara and Banana are distinct tribes and have been said to involve distinct languages (indeed, there are currently three missionary teams, each working on its own Bible translation). In fact these speech forms have about 95 per cent vocabulary in common and very similar grammars. Each is easily intelligible to speakers of the others and they are clearly dialects of one language, on linguistic criteria. Within the Tupí-Guaraní subgroup there is some confusion about what is a language and what is a dialect; the commentary on table 5.1 attempts to provide some clarification.

(b) We have tried to ensure that the standard criteria for proving genetic relationship between languages are followed in this book. The established term for a group of languages that is genetically related is 'language family'. Smaller genetic groups within a family are called 'branches' or 'subgroups' of the family. For instance, one talks of the Germanic branch, or subgroup, within the Indo-European family. Terms such as 'stock' and 'phyllum' (together with 'micro-phyllum', 'macro-phyllum', 'meso-phyllum') go with the quite different methodology of lexicostatistics and are best avoided. We thus talk of the Tupí-Guaraní branch, or subgroup, of the Tupí language family, and so on. (The Portuguese term tronco is sometimes translated
3  GRAMMATICAL TERMINOLOGY

We intend the linguistic surveys contained in this volume to have lasting value. For this reason, among others, we have avoided use of any of the current (or past) formal linguistic theories. The contributors have written in terms of what has recently come to be called Basic Linguistic Theory. This is the cumulated tradition of linguistic description, that has evolved over the last 2,000 years. Most grammars (certainly, all those that have permanent value) are written in terms of the parameters of Basic Linguistic Theory – what is a phoneme; what is a word; the parameters in terms of which systems of tense, aspect and evidentiality vary; the criteria for recognizing a relative clause; and so on.

(a) Contributors have used standard abbreviations for core semantic-syntactic relations:

- **A** transitive subject
- **O** transitive object
- **S** intransitive subject
- **Sₐ** S that is marked in the same way as A on a transitive verb
- **Sₒ** S that is marked in the same way as O on a transitive verb

The terms 'active' and 'stative' are typically used for verbs that take **Sₐ** and **Sₒ** arguments respectively.

A language will thus be characterized as, for instance, AVO, SV (rather than just SVO). There are in fact Amazonian languages where A and S do not occur in the same position, e.g. in Kuikuro, of the Carib family, the basic constituent orders are SV and OVA. (Here a single formula – using 'S' to cover both intransitive and transitive subject – could not satisfactorily be employed.)

(b) Verbs typically fall into a number of subclasses:

- **Intransitive** – can only occur in an intransitive clause, with S core argument.
- **Transitive** – can only occur in an transitive clause, with A and O core arguments.
- **Ambitransitive** – can occur in either an intransitive or a transitive clause; there are two varieties of ambitransitive:
  - **S = A** type, e.g. 'he (S) has eaten', 'he (A) has eaten lunch (O)'
  - **S = O** type, e.g. 'the glass (S) broke', 'he (A) broke the glass (O)'

Most languages have a subtype of transitive:

- **Extended transitive** (or ditransitive) – take A and O arguments and also an oblique argument. Note that, with an extended transitive verb like 'give', in some languages the Gift is in O function with the Recipient in oblique function, in some languages the Recipient is O and the Gift is oblique, while other languages have both construction types available (e.g. in English, *John gave the book to Mary* and *John gave Mary the book*).

A few languages (e.g. Trumai, in §1.2 of chapter 13) also have a subtype of intransitive:

- **Extended intransitive** – take an S argument and also an oblique argument.

(c) Grammatical terms tend to be used in different ways by different authors. It may be useful to provide a characterization of the way some important terms are employed in this volume.

At the clausal level, a language may mark the syntactic functions of its core arguments (A, S, O) either by **dependent-marking**, which involves function markers (case, adpositions, etc.) attached to NPs which realize the core arguments; or by **headmarking**, which involves bound pronominal elements (relating to core arguments) attached to the predicate. (Some languages combine the two strategies.)

**Passive** is a valency-reducing derivation. Prototypically it applies to a transitive clause, taking the A argument out of the core and placing it on the periphery, with the old O becoming new S. Antipassive is similar to passive but here O is moved into the periphery with old A becoming new S. (We know of only one Amazonian language which appears to have an antipassive derivation – Cavineña, in §2.1.2 of chapter 13.)

**Causative** is a valency-increasing derivation. Prototypically it applies to an intransitive clause, bringing in a new argument (the Causer) as A, with the original S becoming O. In some languages causative applies only to intransitives but in others it may also apply to transitives.

**Applicative** is another kind of valency-increasing derivation whose prototypical application is to intransitives. The original S becomes A with what was a peripheral argument being moved into the core as O. There can be a variety of applicatives depending on the original peripheral function of the new O; these include instrumental, comitative, benefactive, dative, locative (see, for example, chapter 3). There is a fuller account (with exemplification) of passive, antipassive, causative and applicative in 'A typology of argument-determined constructions' by R. M. W.
Conventions followed


A serial verb construction involves a single predicate consisting of several verbs which share certain properties – generally, the same subject (S or A) and often other arguments as well. The verbs usually share a single specification for tense, aspect, modality, mood and polarity; they constitute one prosodic phrase.

In some languages certain types of clause combinations (e.g. main clause plus relative clause, main clause plus purposive clause) are obligatorily marked for whether they have the Same Subject (SS) or Different Subjects (DS). This is referred to as switch-reference marking.

Many Amazonian languages have a set of classifiers which characterize a given noun in terms of its inherent properties, typically animacy, shape, form and material. The classifiers may be attached to numerals, deictics, locatives or verbs. Other languages may have a small closed system of genders (or noun classes) which typically include masculine and feminine terms. Each noun must belong to one gender class. A number of Amazonian languages show both classifier sets and gender systems (e.g. Tucano, in chapter 7).

Introduction

R. M. W. DIXON AND ALEXANDRA Y. AIKHENVALD

The Amazon basin is the least known and least understood linguistic region in the world. Maps of the language families of South America (with one colour for each genetic group) purvey an impression of anarchy – there are dabs of yellow and blue and red and orange and brown mingled together like a painting by Jackson Pollock. And when one does get hold of a grammar of an Amazonian language it is likely to show strange properties – multiple sets of classifiers, oddly conditioned ergativity splits, and so on – that constitute exceptions to received ideas about typological universals. In other instances one finds the richest examples of categories that are weakly attested elsewhere. For instance, Tucano languages (chapter 7) have the most highly articulated systems of evidentiality in the world; this is an obligatory specification of the evidence a speaker has for making a statement – whether observed, or reported, or inferred, or assumed. However, a major difficulty is that a high proportion of available grammars are incomplete, affording a glimpse of some exotic grammatical property but with insufficient information to enable the reader to fully understand it, and to realize its overall typological significance.

In fact, the major language families all have markedly discontinuous distributions (more than is found in any other part of the world). The Tupí and Arawak families each comprise at least ten separate geographical regions and the Carib family at least five. Even Tucano, one of the most linguistically homogeneous of families, is spoken over three distinct areas. In addition, there has been a great deal of linguistic diffusion, over large and small regions within the Amazon River Basin. This can make it hard – sometimes impossible – to determine whether points of similarity between two languages in close proximity are evidence of shared genetic development, or simply of mutual borrowing.

One of the editors has devoted several decades to searching for substantive linguistic universals. In case after case, just as he thought he had achieved some significant typological statement, a counter-example popped up; and this was invariably from a language of Amazonia. He decided that the most sensible course of action was to learn Spanish and Portuguese and then go to South America – visit
universities and museums, read the published materials (much of it in obscure jour-
nals), consult unpublished theses, and talk to linguists who have worked on this or
that language (many of whom will never get around to publishing a full grammar).
In this way he achieved a degree of insight into the most complex linguistic area in
the world today.

This book is put forward in a similar spirit. It attempts to describe - in as clear
and accessible a manner as possible - the current linguistic situation in the Amazon
Basin, involving at least 300 languages, belonging to 20 or so language families, plus
more than a dozen genetic isolates. The volume should be regarded as a first and
preliminary step towards elucidating the structural characteristics of Amazonian
languages, and their genetic and areal relationships.

There are several unfortunate aspects to the linguistic situation in Amazonia
today. Only a minority of scholars in linguistics departments in South American
countries work on the indigenous languages, with the remainder preferring to confine
themselves to looking at Spanish or Portuguese in terms of the latest type of formal
theory to filter down from the north. (Seki's Gramática da língua Kamaiurá, cur-
cently in press with Editora da Unicamp, is the first full grammar of a Brazilian lan-
guage, written by a Brazilian, since Anchieta's Arte de grammatica of 1595.) As a
result, a high proportion of the descriptive materials published on South American
languages over the past few decades have been by missionary linguists of the
Summer Institute of Linguistics and similar organizations.

In other parts of the world (for instance, Australia) there is mutual respect and
cooperation between missionary linguists and scholars from the local universities.
In contrast, in most (although not all) South American countries there is antipathy
- sometimes even open hostility - between the two groups. Malicious stories are
spread that the missionaries are agents of the CIA or are engaged in gold explo-
ration or drug trafficking. None of this is true. But many of the missionaries have
had inadequate training and produce 'cookbook' descriptions (in the 1950s and
1960s these were often cast within the impenetrable formalism of tagmemics) that
cannot do justice to the genius of a language. Linguists from universities may
employ other kinds of formalisms, that will soon pass out of fashion. Having made
these general observations, we must add that there are notable exceptions on both
sides - a number of descriptive studies that achieve a high standard of clarity and
explanation.

The Handbook of Amazonian Languages, edited by Desmond C. Derbyshire and
Geoffrey K. Pullum, and published by Mouton de Gruyter, is a most worthwhile
enterprise that has so far run to four volumes that include ten grammars (ranging
in quality from quite good to very good) together with a number of typological and
historical studies. Yet the Handbook would be more useful if its contributors were
not forced to follow an idiosyncratic scheme of organization: syntax, then phonol-
ogy, then morphology. That this is basically unworkable is demonstrated by the fact
that five of the eight grammars in the first three volumes have, as the whole of '23,
Morphology', a single sentence along the lines 'This has been treated in earlier sec-
tions.' It is of course necessary to know the basic inflectional morphology in order
to understand the syntax, so this information is slipped in early on in the descrip-
tion (but at different places in each grammar). The net result may be that nowhere
is there any integrated morphological statement, e.g. of the structure of the verb.

1 CULTURAL BACKGROUND

The Amazon is the world's major river. Its drainage area is over 7 million square
kilometres (more than twice that of the Congo). The volume of water carried is a
fifth of all the flowing water in the world (more than four times that of the Congo).
Its 7 major tributaries are each more than 1,600 kilometres in length (compare this
with the longest river in Europe, the Rhine, which is 1,320 kilometres long. There are distinct wet and dry seasons, at different times of year in different parts of the Amazon Basin, so that the water level may vary by up to 20 metres. Several hundred kilometres up a major tributary, such as the Purús, the river may be about 1 kilometre wide in the dry season but flood to a width of 15 kilometres in the wet season, inundating the adjoining forest lands. North of the Amazon is the Orinoco, flowing into the Caribbean sea, with a drainage area of almost a million square kilometres. In this book we attempt to cover languages spoken in the Amazon and Orinoco Basins – that is, from the north coast of South America, east to the mouth of the Amazon, west to the Andes, and south to the southernmost headwaters of the Amazon tributaries. If most of the languages in a family are spoken in the Amazon/Orinoco Basin (e.g. Arawak) then we cover that family. If most of the languages in a family are outside the region (e.g. Guiicuru) then we do not deal with that family.

Most of the Amazon and Orinoco Basins is lowland rain forest, rich in vegetable and animal foods. Archaeologists (by and large) consider that it was first populated about 12,000 years ago, by hunter/gatherer populations who crossed from Asia at the Bering Strait (which was then dry land) and travelled down the North American coast. It would have taken only a couple of thousand years for Amazonia to be fully populated. The major protein food is fish, supplemented by the hunting of peccaries, deer, tapir, rodents and monkeys. Domestication of plants – focusing first on bitter manioc – is dated to about 5,000 years ago. It is thought to have begun in the higher lands around the foothills of the Andes. People employing slash-and-burn agriculture soon spread out along all the major rivers, while populations of hunters and gatherers contracted to the more remote areas.

Each language family tends to have a characteristic profile in terms of the type of territory it is found in, methods of food procurement, and material culture. Almost all tribes speaking languages belonging to the Arawak, Carib and Tupi families are found in the rain forest, use agriculture, and manufacture canoes, hammocks and pottery. In contrast, Jê-speaking peoples are mostly found on grasslands; they have little agriculture and no canoes, hammocks or pottery; however, they do have the most complex systems of social organization. Scattered between the agricultural tribes, in the heart of the rain forest, are small tribes of hunters and gatherers, belonging to minor linguistic families such as Makú, Mura, Piraẖã and Guahibo. Some of these may be the remainder of earlier populations that occupied larger tracts of land before the agricultural expansion. The languages of the seven Makú tribes, for instance, have been suggested to comprise a rather divergent language family, although full proof has still to be provided for this. They may instead constitute scattered relics of an olden-days linguistic area, whose lan-

Multilingualism was (and is) the norm among the Indian tribes of Amazonia. Some groups – predominantly, those with agriculture – were rather warlike, given to raiding their neighbours. Within some tribes there was a caste of slaves, which originated as captives of war. There were also symbiotic relationships between different groups of people. For instance, the agricultural Tucano tribes (living along the Vaupés River) and the hunting/gathering Makú (living in the forest away from the river) are in a ‘master–underling’ relationship. The Makú supply their ‘masters’ with the meat of forest animals and with fish poison, receiving in exchange manioc flour from Tucano gardens and pottery artifacts. Makú people are considered socially inferior to the Tucano and do not enter into the system of exogamous intermarriage that characterizes the Vaupés region. There are some clans within Tucano tribes (and even one whole tribe) that are said to be of Makú ancestry. Although they now speak a Tucano language and practise agriculture they are considered socially inferior and to be less desirable as marriage partners. (In §1.1 of chapter 3 there is mention of the creation of a mixed language as the result of Carib conquest over an Arawak people.)

Then, in the sixteenth century, came the Europeans. White-skinned invaders quickly took over the coastal areas and major rivers as far as they were navigable. (Chapter 15 explaining how rapids on the Xingu River proved a barrier to the invaders, preserving the Upper Xingu as a refuge area into which Indian tribes congregated.) The Indian population rapidly reduced until, by 1900, it was probably no more than one-tenth of what it had been in 1500.

The biggest killer was involuntary. White invaders brought with them diseases to which Indians had no immunity – smallpox, influenza and the like. European diseases spread ahead of the Europeans themselves; many tribes and languages are likely to have disappeared before even their name could be recorded. Others went a little later. The Arawá people (see chapter 11) are known only from a fifty-word vocabulary collected by the English explorer Chandless in 1867. Almost everyone in the tribe died a few years later from a measles epidemic; the few survivors took refuge with the neighbouring Kulina who are said to have massacred them (Rivet and Tastevin 1938: 72–3).

A typical situation was for a number of tribes, each reduced in numbers, to merge. The ethnologist Franz Caspar (1956: 221) lived for a while among the Tupari tribe and was told by them that ‘in the days of their grandfathers and great-grandfathers several small tribes had merged. Of every man and every woman Tupari was able to say without hesitation of what extraction they were. There was only one man left out of each of the “Vaikorotá”, “Aumeh” and “Mensiátó” tribes. Five were real
“Tupari”. All the rest ... were “Vakarau” ... Even the present language of the tribe, he said, was not the old Tupari, for the minorities had adopted the language of the Vakarau. But people could still recall how the various tribes had spoken and Topo told me a few words of the real Tupari language. Just one of the original languages had survived, almost certainly with considerable substrata from the others.

There was also enslavement by whites. Hemming (1978) tells how the European settlers at the mouth of the Amazon would travel upriver, capture a tribe, bring them back to work on their plantations under harsh conditions, then when they died out (often, within about ten years), go upriver again to capture and enslave another tribe. And there was a great deal of simple murder. A gun is more powerful than bows and arrows; if white people wanted some patch of land they simply took it, together— if necessary — with the lives of the people who had been living there.

Indian tribes invariably fought back; they might win in the short term but never in the long term. Some of them simply fled. At the end of the sixteenth century, eighty-four Tupinambá-speaking villages on the east coast of Brazil became exasperated by the treatment at the hands of the Portuguese and decided to migrate inland. They travelled slowly— up the São Francisco River, across the grasslands of Mato Grosso, until they reached what is now Bolivia, where they were dismayed to encounter Spanish colonists. They then turned north-east and followed the Madeira River until it joined the Amazon, settling on a large island in the middle of the river, which is now called Tumunambara. But there was no escape. In 1639 their island was invaded by Portuguese coming up the Amazon (Hemming 1978: 235).

Before Amazonia was invaded, relations between tribes in Amazonia were sometimes peaceful, other times turbulent. We know, for instance, that speakers of Tupi-Guarani languages spread out over a wide area, with considerable effect on languages of peoples they came into contact with (see §2.2 of chapter 5). There would of course have been instances of migration and of language death. But these happenings greatly increased after 1500. Invasion of the forest hinterlands intensified with the rubber boom at the end of the nineteenth century. This fell away from about 1910 (when cheaper rubber became available from Malaysia), and lands which had been occupied by non-Indians were again released to their original owners. There are, however, rather few Indians remaining, and these are fast being acculturated.

A typical example concerns the Yuqui, speakers of a Tupi-Guarani language. This group lost agriculture and became hunters and gatherers (although they still retained a caste of slaves). For several hundred years they were able to keep out of the way of Spanish colonists, retreating further into the Bolivian forest. By the 1950s the colonial expansion left the Yuqui nowhere to hide. They were in danger of being wiped out when, in the early 1960s, they were persuaded to settle on a New Tribes Mission station (Spearman 1989). Having been saved from one fate, they are now pointed towards another. With their original life-style and culture stripped away (and with only Spanish names), the ethnic identity and language of the Yuqui are unlikely to survive for more than a couple of generations.

Estimates vary, but it is thought that there were, in 1500, somewhere between 2 and 5 million people living in Amazonia. The present Indian population is no more than 400,000. Of the 170 languages reported to be still spoken in Brazil, 115 have less than 1,000 speakers with only 4 having more than 10,000 speakers (none has more than 20,000); similar figures apply for other South American nations. The governments have been slow to extend roads into remote areas or to supply schools and medical posts. But these are coming, and with them assimilation into mainstream society. Every year the indigenous languages are used less and less, and Spanish and Portuguese more and more. Every year another few languages pass into oblivion. Of the estimated 300 languages now spoken in Amazonia, only a small fraction are likely to be still actively used in 100 years’ time.

We should round off this cultural overview with a short comment on trade languages. Several creoles evolved during the early years of colonization, based on Spanish, French and Portuguese. But there was one lingua franca of major importance. What was called Língua Geral ('language that is universal' in Portuguese) or Nheen-gatú ('speech that is good' in Nheengatú itself) evolved on the east coast of Brazil in the sixteenth century. Its morphology was simplified from Tupinambá—a language of the Tupi-Guarani subgroup that was spoken all along the east coast—but the syntax is similar to Portuguese. Língua Geral soon spread up the Amazon and had dialectal variants in different regions. Indeed, it began to compete with Portuguese as the major language of Brazil. In 1727 King John V (back in Lisbon) banned the use of Língua Geral and it began gradually to fade. Língua Geral was the trade language to which most tribes in the Brazilian Amazon were first exposed and it was only replaced by Portuguese in the twentieth century. This creole is not quite extinct, still being spoken as first language by a small number of people in the Upper Rio Negro region.

2 LINGUISTIC DIFFUSION

There are a number of cultural traits that recur throughout Amazonia. These include female initiation rites (which are much commoner than puberty rites for boys). There are generally a number of shamans who control spirits that can both cause and cure diseases. (See, for example, Steward and Faron 1959: 284–318.) The
word for ‘dog’ is often either the same as the word for ‘jaguar’ or else related to it. Probably the most pervasive lexeme is *kuku* or *koko* for ‘mother’s brother’ or ‘father-in-law’ (in a system where one can marry a cross-cousin). If a man marries a woman from another tribe he must know how to address his wife’s father; this has undoubtedly led to the widespread borrowing of the term for ‘wife’s father’.

Linguistic traits, by and large, diffuse more slowly than other cultural traits. However, given sufficient time, languages from several genetic groups that are located in the same geographical area will gradually come to share certain linguistic features and will, as a consequence, make up a ‘linguistic area’. This can be defined as: a region including languages from several different genetic groups, with the languages sharing certain symptomatic features which can be inferred to have diffused across the area. The features will not, as a rule, be found in languages from these genetic groups which are located outside the area. It should be noted that very common properties (those found in very many of the languages of the world, e.g. a tense system) are less significant as diagnostic markers of a linguistic area than are more unusual properties (e.g. an evidentiality system).

Amazonia can be recognized as a linguistic area in terms of features like the following, which are shared by all (or most) languages in the area.

(a) The majority of languages are polysynthetic and head marking; agglutinating with little fusion.
(b) There is typically one liquid phoneme, which is frequently a flap. There are usually more affricates than fricatives. The high unrounded central vowel *i* is frequent. A typical Amazonian vowel system has five members: *i, e, a, i, u/o*. There is typically contrastive nasalization of vowels.
(c) Many languages have extensive classifier and/or gender systems. Gender assignment is often semantically transparent, and is not overtly marked on the head noun.
(d) There are very few oblique cases – often just a locative and an instrumental/comitative.
(e) Possession (either alienable or inalienable) is typically marked on the possessed noun, not on the possessor; the most widespread word order is ‘possessor possessed’ (e.g. ‘John his-canoe’).
(f) Often, just one core argument is cross-referenced on the verb. There may be different bound pronominal paradigms depending on which core argument is being cross-referenced in each particular instance.
(g) The rules for which core argument is cross-referenced can be complex (relating to the meaning of the verb, clause type, etc.) often giving rise to a ‘split-ergative’ system. Fully accusative systems of marking for predicate arguments are rarely encountered.
(h) The bound pronominal forms marking a possessor within an NP are typically the same as one of the bound pronominal paradigms for marking core arguments of a clause (sometimes the same as the A or A/S series, other times the same as the O or O/S series).
(i) Most (although not all) languages have prefixes; there are typically fewer prefix than suffix positions.
(j) If there are several prefix positions, the bound pronominal prefix(es) will typically appear further from the root than prefixes that mark valency-changing derivations (e.g. causative, applicative). (Tucano languages are entirely suffixing; the bound pronominal suffixes appear further from the root than suffixes that mark valency-changing derivations.)
(k) Most verbal categories (e.g. tense, aspect, modality, direction) are expressed through optional suffixes.
(l) Subordinate clauses typically involve nominalized verbs, with the type of subordination being marked on the verb.
(m) If there is noun incorporation, typically only those nouns which are obligatorily possessed can be incorporated, and they typically precede the verb root.
(n) In many languages adverbs and adpositions may be incorporated into the verb, typically following the verb root.
(o) There is generally only a small class of lexical numbers.

There are of course a few exceptions to these pan-Amazonian tendencies. For instance, Tupi-Guarani languages are the only languages in Amazonia which allow incorporation of unpossessed nouns; however, in modern languages this technique is falling out of use (Seki forthcoming, Kakumasu 1986). In Nadèb, a Makú language from the Middle Rio Negro, incorporated adverbs and adpositions precede the verb root instead of following it. In Palikur, a North Arawak language from Brazil and French Guiana, incorporated body parts follow the root (Aikhenvald and Green 1998).

It is interesting to compare typological characteristics of the Amazonian linguistic area, in lowlands South America, with those of the Andean linguistic area in the adjacent mountains, which comprises the Quechua and Aymara families. The Andean area is clearly different in almost all of the characteristics just listed. (b) There are two or three liquids; fricatives rather than affricates; and a three vowel system, *i, a* and *u*, with no contrastive nasalization. (c–e) There are no classifier or...
gender systems; there is an extensive set of case markers; possession is marked both on possessor and on possessed. (f–h) Two core arguments are marked on the verb, in an entirely accusative system; bound pronominal markers of possession show some similarity to, but are not identical with, the forms marking core arguments on the verb. (i–k) There are no prefixes; and there is an obligatory suffixal system for tense and aspect. (l) Subordination does not involve nominalization. (m–n) There is no incorporation of nouns, adverbs or prepositions (Cole 1982: 161 mentions incorporation within nominalization, but this is just a type of compounding, as in English). (o) There is a full set of lexical numbers. As regards characteristic (a), Andean languages are synthetic, and combine head and dependent marking; they are basically agglutinating with some fusion (subject, object and tense suffixes to the verb may be fused).

There is no sharp boundary between the Andean and Amazonian linguistic areas - they tend to flow into each other. For instance, Andean features such as lack of prefixes and an accusative technique for marking syntactic function are found in languages of the Tucano family, which are in Amazonia but fairly close to the Andes.

There are a number of grammatical properties which are not shared by all Amazonian languages but are found in the languages in certain regions and help to define these as linguistic subareas within a wider linguistic area. These include:

1. contrastive lexical tones are found in two regions - one in southern Amazonia (overlapping the states of Rondônia and Mato Grosso) and one in the north-west (along the Vaupés river, from Brazil into Peru).
2. switch-reference marking is found in a group of languages in western Amazonia.
3. gender assignment is not semantically transparent in a region of southern Amazonia centred on the Purús river basin (where Bolivia, Brazil and Peru meet) which includes languages from the Arawá and Chapacura families and the Peruvian and Pre-andine subgroups of Arawak.
4. classifiers are used in different morphosyntactic contexts in different regions - for instance, they are used in possessive constructions in the Upper Rio Negro region but not in Peruvian Arawak languages.

The limited regional distribution of these features can be important for hypothesizing whether certain features were more widespread at some time in the past, and are being progressively lost, or whether they are recent innovations that are currently diffusing more and more widely.

3 GENETIC RELATIONSHIP

There can be a number of explanations for some perceived similarity between two (or more) languages. It can be an indication of genetic relationship. Or it can be the result of areal diffusion, borrowing from one language into another (in one or in both directions) or borrowing by both languages from a third language. Or it can be some universal feature, e.g. the word for 'blow' typically imitates the activity, having a form something like 'pū' (the actual form will relate to the phonetic and phonological resources of each language). Or it can be chance, as where the Australian language Mbabaram and the Indo-European language English both have 'dog' as the name for an animal of the species *Canis*.

To say that a group of languages is genetically related (as a language family) is to say only one thing. It is to assert that they go back to a common ancestor, each having developed from this ‘proto-language’ by its own set of historical changes. Concordant with this, there is only one way to prove that a group of languages is genetically related. This is to propose what the proto-language was like (in some detail) and to describe how each of the modern languages developed, by systematic changes, from this common ancestor.

If two (or more) languages show a certain set of similarities, these need not necessarily be evidence for genetic relationship. They may, alternatively, be due to borrowing, to universal tendencies or just to chance. It is true that certain kinds of similarity (a pronoun paradigm, or a suppletive set) are strongly suggestive of genetic relationship. But suggestion is not proof. And the only proof is to demonstrate the genetic relationship by showing the regular changes which each individual language has undergone from a postulated proto-language (dealing with phonology, verbal and nominal morphology, pronouns, markers of negation, etc., as well as all kinds of lexemes).

In only a few instances do we have written records of sufficient time depth to identify the proto-language and details of the changes modern languages have undergone - Sanskrit and the modern Indic languages; Latin and Romance languages; Old Irish and Irish Gaelic, Scots Gaelic and Manx. In other instances, large portions of a putative proto-language have been reconstructed, together with the changes modern languages have undergone.

A number of proto-languages have been proved in this way, so that all scholars are agreed on their genetic unity; they include Indo-European, Uralic, Dravidian, Austronesian, Algonquian and Mayan. Although most of these language families have been recognized for a fair period of time, it has not proved possible to establish any higher-level genetic linkage between them.
However, some people are not prepared to rest upon what can be scientifically proved and want to go on—to attempt to relate families together as higher-level families, and so on back (sometimes, even back to a putative proto-World language). Such suggestions of higher-level genetic groupings are generally based on a few odd correspondences (often, between one or two languages in family X and a few languages in family Y, rather than between proto-X and proto-Y) and thus cannot be accepted as serious scholarship.

The suggestions that have been made are mutually incompatible, which is itself not a good portent. For instance, Schuller (1919/20) proposed a genetic link of Carib with Arawak (also throwing in Chibcha and Mayan); Greenberg (1960) linked Carib with Jë, Pano and Nambiquara; Greenberg (1987) kept Carib, Jë and Pano together but now placed Nambiquara in a group with Tucano and Makú, among others; and Rodrigues (1985) suggested a genetic link between Carib and Tupí. Thus Carib has been linked with Arawak, Maya, Chibcha, Jë, Pano, Nambiquara and Tupí at one time or another (although not all at once). Kaufman (1990) surveys a number of higher-level genetic suggestions, showing how these disagree; he has lists of languages that are 'Macro-Jë for some [investigators], Macro-Kariban for others', and of languages that are 'Macro-Tupian for some, Macro-Jë for others', etc.

A perceptive comment on suggestions of this type concerning higher-level genetic links comes from Mason (1950: 162): ‘It is a truism of linguistic research that, given large enough vocabularies to compare, and making allowances for all possible changes in the form of a word or stem, as well as in its meaning, a number of apparent similarities, convincing to the uncritical, can be found between any two languages.’ The lack of scientific basis for posited long-distance genetic links can be illustrated with a couple of examples (these could be multiplied a hundred times over).

Ehrenreich (1897) first suggested a genetic link between the Arawá and Arawak families. He gave 17 ‘correspondence sets’ with examples for each drawn from 1 or more of 3 Arawá languages, and 1 or more of 29 Arawak languages. A typical set is:

<table>
<thead>
<tr>
<th>'bow' (as in 'bow and arrow')</th>
<th>Arawá languages</th>
<th>Arawak languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paumari: kudai</td>
<td>Baré: davidaja</td>
<td></td>
</tr>
<tr>
<td>Jamamadi: didiša</td>
<td>Kustenão: turi</td>
<td></td>
</tr>
<tr>
<td>Arawá: biguauha</td>
<td>Tariana: shidoa (‘arrow’)</td>
<td></td>
</tr>
</tbody>
</table>

No proto-form was suggested (nor is it easy to see one that could be suggested), for proto-Arawá, or for proto-Arawak, or for proto-Arawá-Arawak. In fact Arawá lan-

'bow' (as in 'bow and arrow')
Arawá languages | Arawak languages
Paumari: kudai | Baré: davidaja
Jamamadi: didiša | Kustenão: turi
Arawá: biguauha | Tariana: shidoa (‘arrow’)

No proto-form was suggested (nor is it easy to see one that could be suggested), for proto-Arawá, or for proto-Arawak, or for proto-Arawá-Arawak. In fact Arawá lan-

languages all have different (and unrelated) words for 'bow'.

For proto-Arawak, Payne (1991: 396) reconstructs *tapo 'bow' (a form that is in fact reflected in Tariana yawitehephu 'bow'). Yet on the basis of 17 sets of 'correspondences' like this, it was suggested that the Arawá and Arawak families are genetically related. Rivet and Tastevin (1940) followed a similar 'method', drawing odd correspondences between some forms in a few Arawá languages and some forms in a few Arawak languages (and they drew on a pool of 86 Arawak languages). A typical (short) entry from their list of 'cognate' sets is (Rivet and Tastevin 1940: 8):

appeler (‘to call’):

| wawana-he, il appelle [‘he calls’], Kulina [Arawá] |
| pi-úña, bi-úña, Tariana [Arawak] |

The idea that Arawá (now called Madi) and Arawak are related in a family called ‘Arawakan’ was repeated by Matteson (1972). She believed that she was demonstrating genetic relationships by ‘rigorous application of standard techniques of the comparative method’ (Matteson et al. 1972: 21). The first of her 353 putative cognate sets of proto-Arawakan is (Matteson 1972: 172):

1. ‘abdomen’ *tia-ri [in proto-Arawakan]
   *tika-telko, proto-Piro-Apuriná
   *tsa-mon-ii-a, proto-Ashaninka

Note that the only segment that the proposed proto-Arawakan reconstruction and the four intermediate reconstructions have in common is ‘t’.

Further on in the paper, Matteson (1972: 203) gives the Piro and Apuríná forms:

1. ‘abdomen’ *tika-telko [‘t, accent included here but not above]
   [proto-Piro-Apuriná]
   čkête, Piro
   -tkako, Apuriná

and then the forms in her Ashaninka subgroup (Matteson 1972: 211):

1. ‘abdomen’ *tsa-mon-ii-a [proto-Ashaninka]
   -motiš, Machiguenga
   -tsomonté, Campa
   tsomokê, Nomatsiguenga

In the proto-Madi section there is no word for 'abdomen'. However, the first cognate set given is (Matteson 1972: 219):
suggest any genetic connection between South American Indian and Australian Aborigi nal languages (in his 'Comparative study of morphology and syntax in Brazilian Arawakan' (Derbyshire 1986).

There can be a tendency to accept what is put forward as an orthodoxy, without pausing to question it. When a missionary linguist from Peru, who had spent over thirty years studying Kulina (an Arawá language), received a copy of Dixon (1995), she wrote to Dixon: 'I became so intrigued by your declaration of no kinship with the Arawak family that I read and re-read that part [of the paper] ... No wonder we never did seem to “fit the pattern” among Arawaks here.' That is, this linguist had been told that Kulina was an Arawak language. But there are no significant cognates and entirely different grammars, which puzzled her.

We have dwelt on some length upon the Arawak/Arawá suggestion to illustrate its vacuous nature. A similar refutation could be provided for almost all other (perhaps for all other) suggestions of higher-level genetic relationships between the established language families.

There have been a number of commentators who are sound scholars and confine themselves to summarizing the established families and the isolates—as already mentioned, these include Mason (1950), Loukotka (1968), Tovar and De Tovar (1984) and Rodrigues (1986). Rivet has considerable reputation but his work on South American comparative linguistics is deeply flawed (he also suggested a genetic connection between South American Indian and Australian Aboriginal languages). Greenberg's (1960, 1987) 'Amerind' has attracted considerable publicity but is without scholarly foundation (see the critical comments in Dixon 1997: 54-5 and references quoted there).

It should be borne in mind that just because something has been published in a book does not necessarily mean that it has any value. People writing a grammar of language X from family Y sometimes feel impelled to say something like 'family Y belongs to the Macro-Tucanoan stock within the Equatorial-Tucanoan stock within Greenberg's Amerind'. Such a statement should be avoided. It simply adds a veneer of fantasy to what may well be a sound and useful grammar of language X. (Contributors to this volume have confined themselves to just mentioning those genetic groupings that are well accepted and have been proved.)

As a rough rule of thumb, readers should be on their guard against any name beginning with 'macro-' or any unit labelled as a 'stock.' With the possible exception of Macro-Jê (which, if proved to be a genetic group, ought to be labelled) all macro-X's are intensely speculative. A similar comment applies for almost all stocks. (Note that in this book we refer to the Tupi family and the Tupi-Guarani...
branch or subgroup within it – just as one talks of the Indo-European family and the Germanic branch or subgroup within it – in preference to Tupi and Tupi-Guarani family.)

4 THE PUNCTUATED EQUILIBRIUM MODEL

It is generally recognized among historical linguists that one can only recognize historical relationships between languages (and reconstruct shared proto-languages) to a time-depth of 5,000–8,000 years. Given any longer time, individual languages are likely to have changed too much for their original genetic connection to be apparent. Thus, Greenberg’s (1987) views about a relationship between all the languages of South America and most of those of North America are not only unproven but unprovable.

But if one views the ‘family tree’ as the basic model of language relationship, one must presume that there probably was a family tree joining together family trees. That is, that many of the language families (and isolates) of the Americas presumably are genetically related, only one would need a time machine to discover exactly how. In this view, Greenberg’s scheme is a perfectly possible one. But there are many alternative scenarios that are equally plausible, and there is no way of deciding between them.

We believe that this is a mistaken view. The family-tree model is appropriate for certain kinds of language development but not for all of them. Humankind is believed to have had language for at least 100,000 years (many people would opt for a longer time-span). Consider the Indo-European family, which has over 100 modern languages and for which a time-depth of about 6,000 years is posited. There are about 17 periods of 6,000 years in 100,000 years. If one language spawns 10^2 descendants in 6,000 years, then over 100,000 years it should give rise to 10^2 * 10^2 = 10^4 or 10 million billion billion languages. But how many languages do we have in the world today? About 5,000–6,000. This suggests that we should rethink the idea that a family-tree type of language split is the universal model of language development.

Dixon’s essay The rise and fall of languages (1997) puts forward an alternative idea, the Punctuated Equilibrium model of language development. He suggests that during most of the history of humankind there have been long periods of equilibrium, interrupted by short periods of punctuation. During an equilibrium period a given geographical area would have been inhabited by a number of small political groups of similar size, each with its own traditions, religion, laws and language. No one group or language would have substantially greater prestige than any other. Cultural traits, including linguistic features, would steadily diffuse across the geographical area. There would always be gradual changes; some languages would split and some would cease to be spoken, but this would be on a relatively minor scale. In due course, as a result of linguistic diffusion, all the languages in the geographical area would become more similar in structural profile – they would converge towards a common prototype.

An equilibrium situation may be punctuated by some cataclysmic happening. This could be a natural event (flood or drought, rising or falling sea levels); or some material innovation (say, a new method of food production); or the rise of some military or religious leader, intent on building an empire; or entry into new territory. In a period of punctuation there will be expansion and split of peoples and of languages. It is during punctuation that a family-tree model is appropriate; we get the divergence of a steadily increasing set of languages, all emanating from a common proto-language. Eventually, the punctuation mode will gradually lose power, and merge into a new equilibrium period.

When the first people came into South America – 12,000 or so years ago – they would quickly have expanded to fill the continent. Population is likely to roughly double each generation if there is unlimited food and land available. This was a period of punctuation, describable by a family-tree diagram. Once the land was fully populated (which may only have taken 2,000–3,000 years), a period of equilibrium would have commenced within each geographical zone – in the tropical forests, on the grasslands, on the mountains, and so on. There would have been many small groups of hunters and gatherers living in a state of relative equilibrium with each other. Linguistic traits would have diffused across the languages within each region.

A major punctuation would have been triggered by the adoption of agriculture, believed to have taken place about 5,000 years ago. The peoples with agriculture had a distinct advantage in food production. As a result they expanded and split, and so did their languages. This serves to explain the readily provable genetic unity of the Arawak, Carib and Tupi families. The agriculturalists took over the best land, along major rivers. Scattered between them are hunters and gatherers. Genetic relationships are here less clear. As mentioned in §1, it is possible that modern-day Makú tribes are relics of an earlier equilibrium situation, with their similarities being the result of having belonged to an olden-days diffusion area (rather than being indicators of close genetic relationship). The point being made is that there is no sure way to distinguish between similarities that are due to areal diffusion and those that reflect common inheritance. Suppose Europe came to be invaded and settled by the Chinese, leaving just small pockets of people speaking Italian and Basque and Hungarian. A later-day linguist might well take the similarities between these three relic languages (their ‘Standard Average European’ features) as evidence
of genetic relationship. It is possible that the history of the Makú follows a similar pattern.

Chapter 6, on Macro-Jé, discusses the Jé family and 11 other language families, basically all spoken in the grasslands region. Rodrigues presents 39 possible cognates and a number of points of grammatical similarity. All of these are of a typological nature, relating to similarities in construction type, constituent order and grammatical categories between the languages, rather than shared grammatical forms. (Pronominal forms, for instance, appear to rather differ between Macro-Jé families.) Rather than all the Macro-Jé families being related in a higher-level family tree, it seems to us that they could constitute a long-term linguistic area; this would account for their considerable typological similarities. (There may, of course, be genetic links between some of the established families, within the linguistic area.)

It will be seen that — in terms of the Punctuated Equilibrium model of language development — the quest for a family of language families (a tree of trees) is misconceived. Each modern language family probably had its origin in the end of a period of equilibrium. Similarities that are noted between proto-languages may well be areal features which had — during the equilibrium period — spread to all or most of the languages in a given geographical region.

In summary, at the end of the initial human expansion across the whole of South America a family tree would have appropriately mapped the period of punctuation that was drawing to an end. A long period of equilibrium then ensued, with roughly stable population and a great deal of cultural and linguistic diffusion within each ecological zone. The genetic relationships between languages, which had been clear at the end of the period of punctuation, would have gradually become blurred and finally lost as more and more features diffused. Then the equilibrium would have been punctuated and just a few of the languages (those whose peoples had some distinct advantage in living and winning, e.g. agriculture) would expand and split, each starting its own family tree (and obliterating other languages, whose speakers did not have this point of advantage).

Establishing that a group of languages is genetically related, as a language family, is generally an easier matter than deciding on the internal constitution of the family tree, i.e. subgroupings. To assign a set of languages to a subgroup there must be evidence that they have shared some historical development (and it should be some rather distinctive change, not anything that commonly recurs all over the world). The difficulty here is to distinguish between similarities that are due to shared genetic development and those which are due to diffusion. Consider the Arawak family, for instance. A number of subgroups have been tentatively established, each of them being located in a certain region. The languages in a given subgroup do share certain developments. But these may well be characteristic features of that geographical region, and found in both the Arawak and the non-Arawak languages that are spoken there. It is first necessary to examine the areal linguistics of the region, factor out the areal features, and then see whether the remaining similarities between the Arawak languages spoken there constitute sufficient evidence for subgrouping. None of this has yet been done. A full investigation of subgrouping — within an areal perspective — is an important topic for future research.

The European invasion, commencing in 1500, acted as an abrupt punctuation of linguistic areas across all of South America. The prestige languages — Spanish, Portuguese and a few creoles — are continually expanding their domains. It is likely that more than half (perhaps much more than half) of the languages spoken in 1500 have already passed into oblivion, and the remainder are following at a steady rate. Quechua is currently estimated to have about 8.5 million speakers; however, in central Peru (and probably in other regions as well), most children of Quechua-speaking parents are preferring to speak just Spanish (Adelaar 1991: 50). Only Guarani appears to be safe, in the medium term, because it is one of the two national languages of Paraguay (and is in fact spoken in Paraguay by more people than is Spanish). In lowland Amazonia there is little hope for even medium-term survival of any language. The tentacles of European-style civilization have been slow to penetrate the rainforest, and it is this that has helped some cultures and languages to maintain their autonomy. But the outside world is now creeping in. Every decade, each indigenous language is spoken a little less (sometimes, a lot less) and Spanish or Portuguese a little more (or a lot more).

Describing these languages, before they disappear, is an urgent task. If everyone who calls themself a linguist — from South American countries and from overseas — were to devote a year or so to fieldwork, and then write and publish a grammar, dictionary and volume of texts for some previously undescribed (or scarcely described) language, then most of the rich linguistic and cultural heritage would be preserved, for posterity. (This would also lead to substantial enrichment of Basic Linguistic Theory.)

5 ORGANIZATION OF THIS BOOK

Basically, we have devoted a chapter to each of the major language families — a long chapter for a large family (with several dozen languages) and a short chapter for a small family (with just a few members). Tupi-Guarani (a subgroup within the Tupi family) is the best-known group of languages and we have accorded it a chapter of its own.
Then, in chapter 12, Mary Ruth Wise surveys the small language families of Peru. In chapter 13, we provide some information on the small families and language isolates of Brazil, Bolivia, Colombia and Venezuela. The final two chapters look at linguistic areas - the established area in the Icana-Vaupés River Basin, and the incipient area on the Upper Xingu River.

There are a few languages that we say nothing about, for the simple reason that almost nothing is known about them. These include: Awaké, Hoti, Irantxe, Kanoé, Puinave, Supe and the small Katukina family.

Editing the present book has not been an easy task. The contributors come from differing backgrounds and have different kinds of linguistic training. In a couple of instances one of the editors had personal contact with contributors and she was able to work closely with them on their chapters. Chapters 9 and 15 were written in Portuguese and translated into English by the editors. We have not in all cases been able to achieve the degree of coverage that we had hoped for. The reader will be able to see for themself, from this volume, what the state of the art is at the end of the twentieth century - with respect to knowledge of the various language families in Amazonia.

BIBLIOGRAPHY


Loukots, C. 1968. Classification of South American Indian languages. Los Angeles: Latin American Center, UCLA.


INTRODUCTION

The Carib family contains the second largest number of languages (following Arawak) in South America. Hundreds of Carib language names can be found in the literature, but after eliminating different names that apply to only one language or to dialects of a language, the number of languages spoken today is probably about 30. Kaufman (1990: 51) gives the number as 43, but this seems to include some that are probably extinct or are dialects (see Gildea 1998, for the Kaufman 1989 classification). Hoff (1992) gives approximate numbers of speakers of individual languages, and these add up to a total of between 48,000 and 57,000 Carib speakers. Many of them also speak the national language (Dutch, English, French, Portuguese or Spanish).

Carib languages are spoken in six countries of South America: Colombia, Venezuela, Guyana, Suriname, French Guiana and Brazil. Map 1 shows approximate locations for each language. There are differing views about the original Carib homeland. Some scholars claim that it was south of the Amazon, probably between the Xingu and Tapajós rivers. Others locate it north of the Amazon, where most of the Carib groups are currently found, with movements of some groups (including Arara, Bakairi, Kalapalo, Kuikuro and Txikão) to locations south of the Amazon as far as the Xingu river basin. See the introduction in Edwards (1980) for a summary of these views.

The earliest known linguistic materials on any Carib language are two works with the title *Introduction à la langue des Galibis* by the missionary priest Pierre Pelleprat (1606–67). They were based on the research and notes of Fr. Dionisio Mesland, a Jesuit intellectual. The first attempt at a classification came in the next century.

1 It is obvious that this chapter could not have been written without the source materials provided by the scholars whose works are listed in the bibliography. My thanks to all of them for enriching my knowledge of Carib languages and my apologies for any misrepresentations or errors I may have introduced. I am grateful to the following for helpful comments on an initial draft of the paper: Bob Dixon, Spike Gildea, Berend Hoff and Tom Payne. I have especially valued the opportunities to interact in recent years with Gildea and Hoff, both of whom have made such extensive and important contributions to Carib studies.
The best internal classifications in recent times (Girard 1971, Durbin 1977, Kaufman 1989) are reviewed by Gildea (1992, 1998). As he points out, they are all flawed to a greater or lesser degree by lack of adequate data and the consequent impossibility of reconstructing grammatical morphemes and lexical items. Thus Carib comparative and historical studies lag far behind those of the other two large Amazonian language groups, Arawak and Tupí-Guarani. In the last twenty years, however, good descriptions of the morphology and syntax of several Carib languages have become available, and ongoing research by several scholars in other languages should soon lead to more reliable reconstruction and internal classification.

Gildea (1998) is an excellent comparative study of certain features of Carib morphology and syntax (constituent order, person-marking prefixes, tense-aspect-mode marking, marking of possession, nominalizing morphology, and the development of ergativity in main clauses), based on nineteen languages (some not well documented as yet).

The most conservative classification to date is that of Kaufman (1989, reproduced in Gildea 1998), based on cognate retention, consonant cluster reduction, and stress patterns. It is this classification that is reflected in table 2.1, which shows: (1) Kaufman’s putative four main branches of the Carib family (the Carib language and Panare constitute separate groupings); (2) the individual languages on which this survey is based (Kaufman includes others); (3) the two-letter abbreviation used in this paper for each language name; (4) some frequently used alternate names, or the larger units to which they belong; (5) approximate populations (Wai Wai from Hawkins, p.c.; the others from Grimes 1996); and (6) the main sources I have used.

The sources in table 2.1 include substantive grammars for Carib (Hoff 1968), Wai Wai (Hawkins 1998), Hixkaryana (Derbyshire 1979, 1985), Makushi (Williams 1932, Carson 1982, Abbott 1991), Apalai (Koehn and Koehn 1986), and Dekwana (Hall 1988).

There are also useful word lists for Carijona (Otterloo and Peckham ms.), Kashuyana (Derbyshire 1958), Wai Wai (Hawkins, forthcoming), Waimiri-Atroari (Hill and Hill 1994), Makushi (Williams 1932: 138–342), Arekuna and Akawai (Edwards 1980), Dekwana (Hall 1988), Panare (Mättö-Muller and Henley 1990). Other word lists and dictionaries to which I have not had access include: Mosonyi (1978) for Kaljna, a dialect of Carib; Armellada (1943) and Armellada and Salazar (1981) for the Pemong group.

Other Carib languages still spoken, and included in Kaufman’s classification, are:

Yukpa: in a separate Yukpa Group
Japerea: Yukpa Group

made by Fr. Salvatore Filippo Gilij (1721–89), and was restricted to Venezuelan Carib languages. Durbin (1977) supplies information on these early sources. Later classifications appeared, but all were based on insufficient and unreliable data.

2 Carib
2 PHONOLOGY

2.1 Segmental

There is a solid core of segmental phonology that is found in most of the 17 languages listed in table 2.1. I present these segments in tables 2.2 and 2.3. Pending a more reliable reconstruction based on well-established cognate sets, the non-parenthetical consonants and vowels shown in 2.2 and 2.3 may represent something close to a substantial part of the proto-Carib system. The forms in parentheses occur less frequently in some languages, and may prove not to be in the proto-system. They are explained more fully in the notes that follow.

2.2 Phonotactics and suprasegmentals

In most Carib languages, syllable structuring is as follows:

- word- or phrase-initial: (C)(C)(V)(C)
- elsewhere (except AP, see note 3 to table 2.3): C(C)(V)(C)

In the languages with vowel sequences and long vowels (see §2.1 and note 3 to table 2.3), the syllables in which they occur do not have final C, except in DE, which can have a final C or CC following a long (or short) vowel (Hall 1988: 269). In languages in which a syllable-final C occurs, it is usually restricted to a few consonants: in CA, only nasals and h, x, t; in MA, h and n; in WA, p, t, k, m, n, h; in AP, 7 and f; and in PA, 7, h, m, n. Word- or phrase-final C has a very restricted occurrence: none occur in HI other than when apocope of the weak vowel i occurs; only n in TR; m, p, and x in AP; and h in MA (see Durbin 1977: 34). Hoff (1992) gives one additional language, Sikiana (33 speakers). Grimes (1996) gives more details about all of the languages named above, including locations, dialects, alternate names, degrees of bilingualism, and populations.
Table 2.2 (cont.)

Notes (cont.)

MA and DE: alveolar / occurs in CR, AK, AP and BA; palatal / occurs in AK, AP and BA; and the voiced affricate / in CR, AR and AK

(6) The labialized velar / is reported as a distinct segment only in CR. Descriptions of other languages show a two-segment sequence of /kW/; some of these may represent a single /k/ segment.

(7) The voiced stops /, / and / are reported as marginal in some languages (e.g. CA (Hoff 1992), and HI, in which / does not occur at all (Derbyshire 1985)). All three stops occur in CR, AK, BA; only / and / in WA-AT; only palatalized / in KU; only / and / in TX; and / in KA, but only as a non-contrastive variant of /.

(8) The glottal stop is reported for 6 languages: CA, CR, KA, AP, DE and PA; in HI it occurs as a variant of voiceless consonants that precede nasal consonants (/, /, /) or their voiced stop counterparts (/, /, /). WA-AT has pre-glottalized nasals /, / and / which contrast with /, /, and glottalized vowels that contrast with non-glottalized vowels and with each other; there is also non-contrastive glottalization which occurs word finally after short vowels. With all this glottalization going on in Waimiri-Atroari, it is of interest that Hill and Hill (1994) do not show glottal stop as a separate segment in the phonemic inventory, on the grounds that it does not fill a normal C role in the language. It always occurs in conjunction with a vowel or nasal consonant.

(9) In CA, / occurs in word-medial syllable-final position in some dialects and [x] in other dialects, as a result of a historical process of syllable reduction (Hoff 1968, 1992; Gildea 1995a).

Table 2.3 Vowel system found in 17 Carib languages

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>mid</td>
<td>e</td>
<td>(e)</td>
<td>(o)</td>
</tr>
<tr>
<td>low</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

(1) The non-parenthetical segments are found in all 17 languages, with the following exceptions:

i is non-contrastive in HI, occurring only as a variant of /, preceding or following a palatal consonant;

s occurs in some DE examples (Hall 1988), but it is not clear to me from the description of the vowels whether it has phonemic status.

(2) Mid-central vowels are restricted, occurring only as: / in TR, WA, DE and PA; and / in AR and AK.

(3) Long vowels contrast with short vowels in some languages (TR, CR, WW, WA-AT, MA, AR, AK and DE). In other languages there is prosodic lengthening of vowels in CV syllables, predictable in terms of the syllable patterning in words and phrases (CA, HI, in one CA dialect a few suffixes have as part of their form the lengthening of the preceding vowel, and this can perturb the regular prosodic pattern). In AP, nasalized vowels are long in contrast with non-nasalized vowels. Languages that have contrastive long vowels also usually have sequences of two non-identical vowels. AP has sequences of three vowels, with each V constituting a separate syllable (Koehn and Koehn 1986: 122).

Hill 1994), / can be replaced by / or / before a syllable containing /, /, or /.

In KU, vowel harmony is restricted to the vowels of two person-marking prefixes and one possession suffix (Francchetto 1995). For BA, in which voiced consonants are many and occur frequently, Souza (1995) reports a complex system of consonantal harmony, with predictable patterns of alternation, dissimilation and spreading of the voicing feature, affecting lexical items, prefixes and suffixes.

Other assimilatory processes include the following.

In most languages for which morphophonological processes have been reported, there is consonantal fusion (e.g. / + /) and palatalization (e.g. / -> / before /) and / after /). Palatalization and affricativization processes are reported for KU (Francchetto 1995). In WA-AT both palatalization and labialization of / and / occur: syllable-initial / and / can occur as the result of the optional deletion of the vowel / between / and /; and / and / optionally occur when the vowel / immediately follows the initial /; loss of palatalization also occurs in WA-AT: / or / when followed by a /.

2.3 Morphophonology

The following processes are found in Carib languages, usually occurring at morpheme boundaries within a grammatical word: vowel harmony; other assimilatory processes; deletion; insertion; and metathesis.

Vowel harmony across morpheme boundaries is reported for the following languages. In KA, WW and HI / is replaced by / when / occurs in a preceding or following syllable; and in KA and HI, / is replaced by / in prefixes to harmonize with / in a stem-initial syllable. In WA-AT, under certain conditions (listed in Hill and
In CA, a morpheme-final n before p, t or k results in the sequences mb, nd or the velar nasal g respectively. In some CA dialects there has been loss of word-medial syllables which has resulted in breathiness of the vowel in the preceding syllable, which sometimes takes the form of a velar or glottal fricative in syllable-final position (Hoff 1992).

In TR, n > p before k and word-finally; n > m before p; p > φ following h; and k > x following h.

In WW, k > h before t, k, s, tf at morpheme boundaries (cf. HI below); φ > h in consonant clusters and h then deletes if it is preceded by two syllables, a long V, a CC cluster or m.

In HI, morpheme-initial n > p following e, e > j before e, and s and t > f and tf respectively when they follow a syllable containing the vowels e or a; morpheme-final k > h when it precedes a voiceless non-syllabic other than h; and morpheme-initial h > p following morpheme-final m or h. (The last two rules avoid non-permitted CC clusters in a word.)

In WA, p > h before p, m and w; p > m before n; t > h before t, l and n; t > n before m; k > h before k; and k > g before m and n.

In AP, a sequence of V plus nasal C in morpheme-initial position, and preceding another C, becomes a nasalized V; any morpheme-final stop > j before C.

Deletion processes include the following.

High vowels, especially i and u, are deleted in several Carib languages when they occur morpheme-finally following a non-clustered C and preceding a morpheme-initial V or unclustered C, provided that the resulting CC is a permitted one. This process is reported for WW, HI, RA-AT, MA and AP. Whole syllables containing these weak vowels can delete in some languages (CA, MA). For fuller discussion, see Gildea (1995a).

In WW, morpheme-initial y deletes following s and tf. In HI, a similar process occurs: morpheme-initial y deletes when it follows a voiceless alveolar or palatal C, but in this case h is inserted before the C; this results in ty > th, sy > hs, fy > hf (tf and y do not come together at morpheme boundaries).

In MA, certain CV syllables (pi, ti, ki, si) are weak, and when certain affixes are added the whole syllable may be deleted or be reduced to h.

A few insertion processes have been reported.

In WW, when tf is preceded by a long V, the V is shortened and h is inserted following the V.

In HI and MA, there is insertion of j between a prefix that consists of a single V (o or a in HI; u or a in MA) and a stem-initial V (only a V of the same quality as that of the prefix in HI). A similar y-insertion process occurs in PA (Gildea p.c. (See Gildea 1998, for a diachronic hypothesis which traces this y back to a morphological inverse marker in proto-Carib.) In MA, there is also insertion of r between the third-person prefix i- and a stem-initial vowel. In HI, there are also epenthesis rules for the insertion of a brief vowel sound between syllables when certain CC sequences occur.

DE has the following metathesis rules at morpheme boundaries:

\[
\begin{align*}
  i- + t & > ty > tj \\
  i- + k & > ky > tj \\
  i- + s & > sy > s \\
  i- + m & > my > n \\
  i- + n & > ny > n \\
  i- + r & > ry > y
\end{align*}
\]

HI has one invariable metathesis rule: morpheme-final s or f followed by h changes to fs and hf respectively (cf. the above deletion of y process, which also entails insertion of h). There are other more idiosyncratic changes, where certain CV sequences become VC when followed by certain suffixes or particles: hu>uh, hi>ih, ti>ht. MA has one similar process that regularly occurs at morpheme boundaries: hi and hu>ih and uh respectively.

3 Morphology, particles and pronouns

This section describes: inflectional morphology (§3.1), derivational morphology (§3.2), the particle word class (§3.3) and pronouns (§3.4).

Verbs in Carib languages are especially complex morphologically, and nominals and adverbials also abound in prefixes and suffixes, both derivational and inflectional. The structure of the verb in most Carib languages is:

Person prefix – de-transitivizer prefix – root – derivational suffix – tense/aspect/mood/number suffix

The ergative languages have an A person marker attached to an ergative marker (§3.1.1). For MA, Abbott (1991) treats the whole sequence as a suffix of the verb, in final position following the TAMN suffix. For KU, Franchetto (1990) treats the sequence as a separate word, immediately following the verb.

3.1 Inflectional morphology

There are four main types of inflectional morphology discussed in this subsection: person-marking affixes on verbs and nouns, and on some adverbials and postpositions (§3.1.1); tense, aspect, mode and number suffixes on verbs (§3.1.2); possession, tense and number suffixes on nouns (§3.1.3); and inflectional suffixes on locative postpositions (§3.1.4).
3.1.1 Person-marking affixes on verbs, nouns, adverbials and postpositions

In order to get a clear picture of the person-marking systems of Carib languages, it is necessary to distinguish the three grammatical functions to which the verb person markers relate: intransitive subject (S), transitive subject (A) and object (O). There is only one prefix per verb.

A characteristic of the morphosyntax of these languages is that they show varying degrees of ergatively organized morphology and syntax; in some, ergativity is strong, in others, it is weaker. (See §4.3 for further discussion.) The majority of the languages have split systems, partly ergative and partly accusative. The person-marking systems form four distinct sets that can be classified as follows:

- Transitive A-oriented prefixes
- Transitive O-oriented prefixes
- Intransitive S_{o}-oriented prefixes
- Intransitive S_{a}-oriented prefixes

(See Derbyshire 1991 and 1994 for the HI system based on this classification; and Gildea 1998, for an alternative analysis which includes an inverse system and a summary of other possible analyses.)

S_{o} prefixes occur with one set of intransitive verbs (in general, 'active' verbs), while the S_{a} prefixes occur with another set (in general, 'stative' verbs). However, the semantic distinction between 'active' and 'stative' does not always determine to which set a particular verb belongs. Some languages listed in table 2.5 have only a partial set of S_{a}-oriented forms (restricted to one or two persons). In PA the distinction is not primarily between S_{a} and S_{o}, but between two tense/aspect paradigms: 'past perfective', which takes the A or S_{a} prefixes, and 'non-trans-perfective', which takes the prefixes: -1'S_{o}, -2'S_{o}, and -3'S_{o} (only the first and second person forms are the same as those in the O set shown in table 2.5; see Gildea 1998).

Gildea (1998) has reconstructed the proto-Carib forms of the person-marking prefixes based on cognate sets from nineteen languages. In tables 2.4 and 2.5, I use his reconstructed forms, but have modified the cognate sets to conform to the data in my own sources and to include only the languages on which this study is based. I exclude AK, AR, MA and KU, the strongly ergative languages, whose person-marking systems are somewhat different (see table 2.6 for the MA and KU sets), and also two languages for which my sources do not provide the necessary data - KA and TX. There is one category of person not shown in the tables: first person exclusive (1 + 3); except in MA and KU (see table 2.6), the 1 + 3 prefix is identical in form and function with third person, and a free pronoun ana (or cognate) '1 + 3' is always present. In PA, the 1 + 2 prefix (n(i))- is also identical with third person.

Table 2.4 Comparative set of transitive A-oriented prefixes which occur when the O is third person

<table>
<thead>
<tr>
<th>Language</th>
<th>1A</th>
<th>2A</th>
<th>1+2A</th>
<th>3A</th>
</tr>
</thead>
<tbody>
<tr>
<td>proto</td>
<td>*ti(1)-</td>
<td>*mi(1)-</td>
<td>*kiti(1)-&gt;</td>
<td>*n(i)-</td>
</tr>
<tr>
<td>AP</td>
<td>i-/o-</td>
<td>m(i)-</td>
<td>s(i)-</td>
<td>n(i)-/kin(i)-</td>
</tr>
<tr>
<td>BA</td>
<td>s-/k-</td>
<td>m-/m-</td>
<td>ki-/kize-</td>
<td>n-/</td>
</tr>
<tr>
<td>CA</td>
<td>s(i)-</td>
<td>m(i)-</td>
<td>ki(i)-</td>
<td>n(i)-</td>
</tr>
<tr>
<td>CR</td>
<td>i-</td>
<td>m(i)-</td>
<td>kit(i)-</td>
<td>n(i)-</td>
</tr>
<tr>
<td>DE</td>
<td>w-</td>
<td>m-</td>
<td>k-</td>
<td>n-</td>
</tr>
<tr>
<td>HI</td>
<td>w-/i-</td>
<td>m(i)-</td>
<td>t(i)-</td>
<td>n(i)-</td>
</tr>
<tr>
<td>PA</td>
<td>t(i)-</td>
<td>m(i)-</td>
<td>n(i)-</td>
<td>n(i)-</td>
</tr>
<tr>
<td>TR</td>
<td>w(i)-/i-</td>
<td>m(i)-</td>
<td>k(i)-/kit</td>
<td>n/-kin-</td>
</tr>
<tr>
<td>WW</td>
<td>w(i)-/o-</td>
<td>m(i)-</td>
<td>t(i)-/th(i)-</td>
<td>n(i)-/o-/p-</td>
</tr>
<tr>
<td>WA</td>
<td>w(i)-/o-</td>
<td>m(i)-</td>
<td>k-/kat/h-/s-</td>
<td>n(i)-/n(i)-mën-/k(i)-/-o-</td>
</tr>
</tbody>
</table>

Notes:
The BA prefix -'1/30' is presented here as A-oriented, i.e., with the S referring to '1A'; some linguists present it as O-oriented, i.e., with the S referring to '3O' (e.g. Souza 1994: 23-4: 'the patient [=O] is always the term referred to' [my translation]). Souza, however, makes exceptions for the 2A/3O and 1+2A/3O prefixes, which refer to the A ('subject') is the term she uses). Pending a fuller description of BA, I tentatively place -'1/30' in the A-oriented set.

(both in the 1+2A/3O prefix (table 2.4) and the 3A/1+2O prefix (table 2.5), thus collapsing the distinction between 1+2 and 1+3 in the verb prefix, retaining it only in the free pronouns (table 2.11)).

The intransitive verb S_{o} prefixes are identical, or very nearly so, with most of the forms in table 2.4. The only significant differences are with the first person forms in five of the languages: BA i-/o- '1S_{o}' (cf. s-/k- '1A'; according to Souza (1994) there is no S_{o}/S_{a} distinction in BA); CA a- '1S_{o}' (cf. s(i)- '1A'); WW k(i)- '1S_{o}' (cf. w(i)-/o- '1A'); HI k(i)-/h- '1S_{o}' (cf. w(i)-/o- '1A'); PA in(i)-/a- '1S_{o}' (cf. in(i)- '1A'). In PA there is also a suffix, -në '1+2S' (cognate with the MA suffix -ni '1+2S' in table 2.6), which occurs in a subset of main clauses that are part of the reanalysis hypothesis of Gildea (1998); see also §4.3.

The absence of third-person prefix in table 2.5 requires explanation. In most of these languages the O-oriented prefixes function as part of a system governed by a person/agentivity hierarchy: first person and second person outrank third person; when the higher-ranked person in a transitive clause is the subject, the A-oriented prefix is used, marking the A; when the higher-ranked person is the
Table 2.5 Comparative set of transitive O-oriented prefixes which occur when the A is third person

<table>
<thead>
<tr>
<th>Language</th>
<th>1O</th>
<th>2O</th>
<th>1+2O</th>
</tr>
</thead>
<tbody>
<tr>
<td>proto</td>
<td>*(y)-</td>
<td>*(y)-</td>
<td>*(i)-</td>
</tr>
<tr>
<td>AP</td>
<td>y(i)-</td>
<td>o-</td>
<td>(k)-</td>
</tr>
<tr>
<td>BA</td>
<td>i-</td>
<td>o-</td>
<td>i-</td>
</tr>
<tr>
<td>CA</td>
<td>a(y)-</td>
<td>a(y)-o-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>CR</td>
<td>y(i)-</td>
<td>e(i)-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>DE</td>
<td>o(y)-</td>
<td>(o(d))-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>HI</td>
<td>r(i)-</td>
<td>o(i)-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>PA</td>
<td>u(y)-</td>
<td>o(i)-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>TR</td>
<td>y(i)-y(i)</td>
<td>a(i)-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>WW</td>
<td>o(y)-</td>
<td>a(w)-</td>
<td>k(i)-</td>
</tr>
<tr>
<td>WA</td>
<td>y-</td>
<td>o(w)-</td>
<td>k(i)-</td>
</tr>
</tbody>
</table>

The O-oriented prefix is used, marking the O. The O-oriented set occurs when the subject of a transitive verb is '3' and the object is '1' or '2' or '1 + 2'. When only first and second persons are involved (1A/2O or 2A/1O), the person markers vary among the languages, and usually one of the participants is expressed with a free pronoun. (This is not, however, the case in CA, where both 1A/2O and 2A/1O have the same prefixal form, and use of a free pronoun does not disambiguate; only situation and context can do so.) When both subject and object are third person, neither outranks the other; the prefix then used is from the A-oriented set. Some features of these O-oriented transitive sets have been explained in terms of an inverse system (Doris L. Payne 1993; Gildea 1994, 1998).

The intransitive S_0 prefixes are identical, or nearly so, with the O forms in table 2.5 in four languages: CA, DE, TR and WA. In two languages, AP and HI, S_0 prefixes occur only for second person and the forms are identical with the O forms for that person; S_0 prefixes occur for all other persons. In two languages, CR and WW, there is no S_0 set; all intransitive verbs take the same prefixes, which in most cases are from the A/S_0 set; in CR, however, the y(i)- '1S' prefix is from the O/S_0 set. In the remaining languages information is lacking as to whether there is a distinct intransitive S_0 set or subset of prefixes.

The following examples (1a-1d) from Dekwana illustrate the person-marking prefixes on transitive and intransitive verbs (taken from Hall 1988):

(1a) Transitive A-oriented: w-edant(0)-a
    1A/3O-meet-PRES
    'I meet him/her.'

(1b) Intransitive A-oriented: w-amo-a
    1S-cry-PRES
    'I am crying.'

(1c) Intransitive S_0-oriented: w-amo-a
    1S-cry-PRES
    'I am crying.'

(1d) Intransitive S_0-oriented: y-a:wo-a
    1S-swell-PRES
    'I am swelling.'

The strongly ergative languages have different person-marking systems, although the forms are clearly cognate to the forms found in table 2.5, which shows the O-oriented set of the mixed ergative-accusative languages. The four ergative languages fall into two groups geographically: AR, AK and MA, which are located north of the Amazon in Guyana and Brazil; and KU, located in the Xingu River basin well south of the Amazon. MA and KU are shown as representative of the two groups in table 2.6.

The most striking difference between these two sets and the A- and O-oriented sets in tables 2.4 and 2.5 is that the (ergative) A forms are suffixes (MA), or are postposed to the verb as part of a separate word, composed of the pronominal proclitic and the ergative enclitic (KU). In general, the same forms are used for marking both the A (suffix or clitic) and S/O arguments.

In MA, but not in KU, the 1 + 2S form -n/i) is unusual in two ways: it is a suffix, and its form is totally different from the corresponding 1 + 2O prefix, and also it does not follow the normal Carib pattern for 1 + 2S. (Cf. the PA suffix -ne '1 + 2S' described earlier.) There are also different forms in MA for 2S/O (a(w)-a(y)-prefix) and 2A (-o-ya suffix).
The A forms in table 2.6 include the ergative clitic -ya in MA, and -héke in KU. (In other Carib languages -ya is cognate with forms like the HI postposition -wya ‘to’ (indirect object marker) or ‘ergative/agentive’; -héke is probably cognate with forms like the HI postposition hoko ‘concerning, about, with’.) These ergative clitics occur on the free form noun when there is one, or are attached to the pronominal suffix (MA) or proclitic (KU), as described above. In MA, the free form and person-marker never co-occur. The forms are clearly not cognates with each other; the MA -ya, however, is cognate with clitics that sometimes occur with an ergative meaning in the languages listed in tables 2.4 and 2.5. In MA the -ya clitic does not occur for ‘1 + 2’ person, but -héke does occur in KU, in the form ku-péke. KU is unusual among Carib languages in having a ‘1 + 3’ free form pronoun, tísu, that is not a cognate of anna (see table 2.11), and whose corresponding prefix (tí-) is not the third person. The person/agentivity hierarchy is not applicable to the person-marking in the ergatively organized systems of these languages, but in KU it is relevant to what Franchetto (1990) calls the ‘interactive moods’, in which ergative constructions occur with third person A, and accusative constructions occur with first person and first person inclusive A (with first person exclusive and second person A there is free alternation between the ergative and accusative patterns). See §4.3 for further discussion on ergativity.

Examples of person markers in Makushi (2a–c) and Kuikuro (3a–c) are:

(2a) i-koneka-hpi-u-ya
30-make-PAST-1A-ERG
‘I made it.’

(2b) u-koneka-hpi-i-ya
10-make-PAST-3A-ERG
‘He made me.’

(2c) u-komami-hpi
1S-stay-PAST
‘I stayed.’

(3a) u-ta-li-íg i-héke
10-hear-PUNCT-FUT ASPECT 3A-ERG
‘He will hear me.’

(3b) u-katfun-tari
1S-work-SUBJ
‘I am working.’

Possessed nouns, certain derived adverbials, and some postpositions (mostly locatives) have prefixes that refer to the possessor (which is the underlying S or O in forms derived from verbs; see §3.2). The forms are nearly identical with the O-oriented prefixes in table 2.5, for the languages listed there, and the S/O prefixes in table 2.6, for the ergative languages. The main difference is the addition of a third-person prefix, which does not occur in the O-oriented set, for reasons explained earlier in this section; the form of the prefix is i-, í-, ð- ‘3POSSR’ in most languages (reconstructed as proto-Carib *y- by Gildea 1998). There is also a third-person reflexive possessor form, reconstructed by Gildea as *t- ‘3REFL’, and having that same form in nearly all languages. See §3.1.3 for possession, tense and number suffixes on possessed nouns, and §3.2 for derivational affixes which form nouns and adverbs. Examples from Hixkaryana, a mixed ergative-accusative language (4a–b) and Makushi, an ergative language (5), are:

(4a) ay-amori o-amori i-hanari t-hanari
2-hand 3-hand 3-ear 3RELF-ear
‘your hand’ ‘his hand’ ‘his ear’ ‘his own ear’

(4b) ro-hyawo i-hyawo tâ-hyawo
1-with 3-with 3RELF-with
‘with me’ ‘with him’ ‘with himself’

(5) u-wahkari i-wahkari ti-pon i-pikiri
1-axe 3-axe 3REFL-clothes 3-behind
‘my axe’ ‘his axe’ ‘her own clothes’ ‘behind him’

3.1.2 Tense, aspect, mode and number suffixes on verbs
Verb suffixes in the indicative mood are portmanteau forms expressing tense, aspect, mode and number. (PA, and possibly some other Carib languages, do not have the number component.) There is another set of suffixes for the imperative (including hortatory) mood; these have components of number and motion, and differ in form according to the person of the subject.

Here I will describe only the indicative TAMN suffixes. The mixed ergative-accusative languages for which TAMN data are available vary widely as to the number, tense and aspect distinctions that are made. Hoff (1968) shows nine for CA. Gildea (1998) reports three past tenses with no aspectual distinction for PA.
Table 2.7  Some tense-aspect suffixes in mixed ERG-ACC languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Past or immediate past</th>
<th>Recent past</th>
<th>Distant past</th>
<th>Nonpast</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>-no</td>
<td>-nel-sene</td>
<td>-V(ana)ko</td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>-da</td>
<td>-ma,-wa</td>
<td>-da,-raki,-taki</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-i (?)</td>
<td>-yakon</td>
<td>-ya,-saa,-e</td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td>-o</td>
<td>-ne</td>
<td>-ya,-yanе</td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>-i,-ano</td>
<td>-ne,-akene</td>
<td>-a,-ta,-tai</td>
<td></td>
</tr>
<tr>
<td>HI</td>
<td>-no</td>
<td>-yakon/-yakoni</td>
<td>-yaa,-sha,-yanо</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>-yah</td>
<td>-i,-e</td>
<td>-ya(ke)</td>
<td>-a,-pe</td>
</tr>
<tr>
<td>TR</td>
<td>-o</td>
<td>-ne,-o</td>
<td>-ya(-e)</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>-o,-o,-u</td>
<td>-ye,-y,-yakpe,-ekpe</td>
<td>(ya)i,-(a)i,-(e)i</td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>-o</td>
<td>-ne</td>
<td>-ya,-i</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
(1) The distinction between recent and distant past is neutralized in DE, WW and WA. I am assuming this is also the case in BA, CA and CR, but the data in my sources may not be complete.
(2) In DE, in the two past tenses, the first forms -í and -/le also mark completive aspect, and the other two forms, -ano and -ake/-le, mark incomplete.
(3) The completive vs aspectual feature aspect distinction for the recent and distant past tense forms in AP and HI is indicated by the slash (/).
(4) In CA, the form shown for immediate past, -i, does not have as its primary value ‘past’ but a modal value ‘realis’ (Hoff 1968: 172-3, 177-9; 1986: 81-6).

Hawkins (1998) describes three past and two nonpast tenses, with no aspectual distinctions, for WW. Derbyshire (1985) records five past and two nonpast tense paradigms for HI; there are three basic past tenses, the recent past and distant past having aspecual distinctions; the two nonpast tenses have a modal distinction.

I present the suffixes found in my sources in table 2.7. These are part of Gildea’s (1998) Set 1 TAM(N) suffixes, except for the AP nonpast form, which Gildea considers an innovation. His Set 2 suffixes, which occur in a few Carib languages, are also innovations, resulting in most cases from the reanalysis of some of the nominalizing suffixes (see table 2.10; and discussion in §4.3). I suspend final judgement on the status of Gildea’s Set 2 suffixes until Carib languages have been sufficiently documented so that his reanalysis hypotheses can be thoroughly tested.

The collective number component of the TAMN suffixes, which refers to the S, A and/or O referents, constitutes a cognate set in all the languages for which data are available: -to (AP, CA, DE); -ti (HI); -tu (WW); -te, -teo (WA). Some of the -ya(C)V forms in table 2.7 are morphologically complex, as can be seen from the placement of the number suffix; cf. the following noncollective (6a, c) and collective (6b, d) HI verb forms:

(6a)  n-amæk-yako
3-hunt-REC.PAST + COMPL
‘He went hunting.’

(6b)  n-amæk-ya-(fo)-ko
3-hunt-REC.PAST-COLL-REC.PAST + COMPL
‘They went hunting.’

(6c)  k-omok-yakoni
1-come-DIST.PAST + CONT
‘I used to come.’

(6d)  m-omok-ya-(j)-konji
2-come-DIST.PAST,COLL-DIST.PAST + CONT
‘You all used to come.’

Sometimes the collective suffix (6f) replaces the noncollective form (6e) and constitutes the complete tense-aspect marker, again illustrated from HI:

(6e)  o-wayeh-no
3-die-IMM.PAST
‘He died.’

(6f)  o-wayeh-(fo)-wi
3-die-IMM.PAST + COLL
‘They died.’

The ergative languages vary significantly in respect of both the TAMN system (fewer and different categories) and the forms of the suffixes (or absence thereof). MA has three tense distinctions, analysed by Abbott (1991) as: ‘past’ (previous day or earlier); ‘proximate’ (indicated with an auxiliary, not a suffix, and referring to a future); and ‘universal’ (no overt form, and referring to present or future). There is also a distinct perfective/perfect aspect which can occur alone or co-occur with the past (which occurs on an auxiliary) and proximate tenses. The copula and two auxiliary verbs function in the MA TAMN system. (Other MA aspect markers appear to be outside the system; see Abbott 1991.)

KU has at least three tense-aspect distinctions: continuative, punctiliar and future (the latter two can co-occur). There is another clitic, -leha ‘aspect’, without
further specification. There is also a suffix which functions as both nominalizer and perfective aspect (Franchetto 1990). The forms in these two languages are:

<table>
<thead>
<tr>
<th>Makushi</th>
<th>Kuikuro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past</td>
<td>Continuative</td>
</tr>
<tr>
<td>Proximate</td>
<td>Verb + pih + copula</td>
</tr>
<tr>
<td>Universal</td>
<td>Uninflected verb</td>
</tr>
<tr>
<td>Perfective/Perfect</td>
<td>-sah</td>
</tr>
</tbody>
</table>

The collective suffixes are: KU -ko; MA -kon, -nikon. Examples showing full intransitive verb forros with only noncollective (7a, 8a) and collective (7b, 8b) suffixes (where (7) is Makushi and (8) is Kuikuro) are:

- **Makushi**
  - (7a) -pih
  - 3S-at 3S-eat-PAST
  - 'He ate him.'
  - (7b) -pih-kon
  - 2S-eat-PAST-COLL
  - 'You all ate.'
  - (8a) -pih
  - hard 3-work-CONT
  - 'He is working hard.'
  - (8b) -pih-ko
  - hard 3-work-CONT-COLL
  - 'They are working hard.'

3.1.3 Possession, tense and number suffixes on nouns
Possession parameters include inalienable/alienable and present/past time. Noun suffixes indicate possession and time. The person of the possessor is indicated by a prefix (§3.1.1). Collective number is marked by a postposed particle (HI kono) for simple, non-derived nouns and also for many derived nouns; other HI derivations take the suffix -yamo 'COLL'. (See §3.2.2 for examples of the derivational morphology.) I will first illustrate the nominal inflection system from HI, and then note some significant differences that have been reported for other Carib languages.

a) Hixkaryana nominal inflection system
The following suffixes occur to mark possession in HI. The basic forms are lexically conditioned, but most possessed noun stems occur with -ri or -ni:

- **Basic possession:** -ni, -ti, -tfe, -a
- **Past/former possession:** -thiri/thiriri, -thritho 'simple past/former possession'; and -nhirihno 'remote past/former possession' (the alternate forms in the first two pairs are phonologically conditioned; the -tho and -nho forms occur with the prefixes for first person, first person exclusive, and third person with a preceding NP; the -thiri and -nhiri forms occur with the other person-marking prefixes).

Inalienably possessed nouns (mainly body parts and kinship terms, but also a few others) normally have the possession suffix. (There is a suffix -nano 'depossession' which replaces the possession suffix on inalienably possessed nouns, and gives the item more general reference, without specifying any possessor, but it is rarely used in HI, except for one or two items. Cognates of this depossession suffix occur in other Carib languages.) The possession suffix is obligatory (and always the -ri form) in three HI derivational processes (see §3.2.2). Alienably possessed nouns usually occur with the -ni form when a possessed form is required, but -ri occurs with some nouns, including kanawa 'canoe' (9), which is alienably possessed; the others in (9) are inalienably possessed.

HI non-derived possessed nouns:

- (9) i-kanawa-ri ro-yo-ni k-osoti  o-he-tfe
  - 3-canoe-PossD 1-mother-PossD 1 + 2-name-PossD 2-wife-PossD
  - 'his canoe' 'my mother' 'our (two) names' 'your wife'

Compare the use of the collective particle: ikanawari kono 'their canoes'; kosot kono 'our (three or more) names'.

The suffixes occur when there is a free form possessor: romura kanawari 'my son's canoe'; rowi yosoti 'my brother's name'. The only prefix that co-occurs with a free form possessor is y- 'GEN' when the possessed noun is vowel-initial, as in y-osoti (GEN-name) in the preceding example.

HI past/former possession: kokanawatho 'my old (or former) canoe'; ikanawarikiri 'his old/former canoe'; ohetfenkiri 'your former wife'; romura hetfenho 'my son's former wife'.

b) Variations in other Carib languages
Some languages have a very similar set of suffixes to HI for marking basic possession (AP, CA, WW). The past/former possession suffixes in two of these languages are: AP -tiri or -vimiri; WW -thiri, -tho, -nhiri, -nho (the same forms as HI, except for the HI initial t variant). In CA, the suffix sequence -ri-mbo (-POSS-D-PAST) corresponds most closely to the HI, AP and WW forms, but -mbo is not restricted to possessive constructions, and its semantic reach is wider than 'past' (Hoff p.c.).
DE suffixes are -dû, -i and a few less common variants; the -dû is probably a cognate of -ri which occurs in other languages. There are three collective suffixes and these replace the noncollective suffix: -tomo marks plurality of the entity referred to by the noun; -komo marks plurality of the person of the possessor; and the sequence of the two takes the form of -tonkomo. There is also a past suffix, -hû, which co-occurs with, and precedes, the possession suffix -dû.

TR possessed nouns usually have no overt suffix, although -ri optionally occurs under certain conditions; the person prefixes are the primary indicators of possession; the past/former possession suffix is -hrpê; collective suffixes are -kon-komo and -ton-tomo.

WA possession suffixes are -n, -i, the collective suffix is -kom; past/former possession suffixes are: -(V)rpê, -(V)npê, -(V)pi, -(V)pi.

The ergative language MA has the possession suffixes -i, -ni, -ka, -no. The collective suffix is -kon. The past/former possession suffix is -hpi on non-derived nouns and -hpi (the verb past-tense suffix) on derived nouns.

There are many obligatorily non-possessed nouns in Carib languages, including all names of plants, animals, persons and natural phenomena. For some of these a generic noun that has possession marking is often used as a classifier, with or without the addition of the specific noun, e.g., HI rokni (okayisu) 'my pet (a dog)' (rokni I-domicile animal-POSSD, okayisu 'dog'). See Koehn (1994) on AP genitive constructions; and Carlson and Payne (1989), who first described this system as 'genitive classifier', established a typology of such systems, and provided examples from AP, MA and PA. Carib languages tend to have between ten and twenty of these classifiers, and it is the only type of noun classifier system found in the family.

3.1.4 Inflectional suffixes on locative postpositions

At the end of §3.1.1 I noted that locative postpositions could take person-marking prefixes. In Carib languages there is a set of locative suffixes that occur with simple, one-syllable postposition stems, and also on more complex locative stems. These carry basic locative meanings such as 'in/on', 'to/into', 'from', 'passing by/through', and the forms vary according to the meanings of the nouns which they govern. The primary semantic distinctions are usually: liquid (e.g. water), flat surface, open area, enclosed place. More specific locative meanings result when the basic suffixes are attached to complex stems. Table 2.8 shows part of a basic set from each of four languages (AP, HI, MA, WW).

The distinction between 'open area' and 'enclosed place' is somewhat fuzzy. It is only specific for MA (Abbott 1991). The general pattern, however, is clear. The final suffixal forms (-na, -e, -o, -ka, etc.) constitute part of the basic set of locative suffixes (there are a few other basic forms); and they are cognates in three of the languages (MA, the ergative language, being the exception in some of its forms). The preceding forms (the locative stems) reflect the four primary semantic distinctions, and these seem to be cognates in all four languages. More complex locative stems occur with the same final suffixes, and amplify the primary set of semantic distinctions. These facts can be seen in the following HI examples (Derbyshire 1985: 205–19 lists twenty-one semantically conditioned sets in HI):

(10a) asama y-aheta-wo
trail GEN-edge-at
'at the edge of the trail'

(10b) wewe mahya-ye
tree rear-from
'from behind the tree'

3.2 Derivational morphology

Carib languages are rich in derivational morphology: verb roots derived from nouns and from verb roots; compounds of verb and noun roots; nouns derived from

<table>
<thead>
<tr>
<th>Liquid</th>
<th>Apalai</th>
<th>Hixkaryana</th>
<th>Makushi</th>
<th>Wai Wai</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>kua-o</td>
<td>kwa-w</td>
<td>ka</td>
<td>kwa-w</td>
</tr>
<tr>
<td>into</td>
<td>kua-ka</td>
<td>kwa-ka</td>
<td>ka-ta</td>
<td>kwa-ka</td>
</tr>
<tr>
<td>from</td>
<td>kua-e</td>
<td>kwa-ye</td>
<td>ka-pai</td>
<td>kwa-y</td>
</tr>
<tr>
<td>Flat surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>out</td>
<td>po</td>
<td>ho</td>
<td>po</td>
<td>φo</td>
</tr>
<tr>
<td>toffonto</td>
<td>po-na</td>
<td>ho-na</td>
<td>po-na</td>
<td>φo-na</td>
</tr>
<tr>
<td>from</td>
<td>po-e</td>
<td>ho-ye</td>
<td>po-i</td>
<td>φo-y</td>
</tr>
<tr>
<td>Open area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in/on/outft</td>
<td>ta-o</td>
<td>ta-woly-wa</td>
<td>ya</td>
<td>ya-w</td>
</tr>
<tr>
<td>toffinto</td>
<td>ta-ka</td>
<td>ta-ka/ya-ka</td>
<td>ya-pih</td>
<td>ya-ka</td>
</tr>
<tr>
<td>from</td>
<td>ta-e</td>
<td>ta-ye/-ya-ye</td>
<td>ya-pai</td>
<td>ya-y</td>
</tr>
<tr>
<td>Enclosed place</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in</td>
<td>e-o</td>
<td>ya-wo</td>
<td>ta</td>
<td>ya-w</td>
</tr>
<tr>
<td>to</td>
<td>a-ka</td>
<td>ya-ka</td>
<td>ta-pih</td>
<td>ya-ka</td>
</tr>
<tr>
<td>from</td>
<td>e-e</td>
<td>ya-ye</td>
<td>ta-pai</td>
<td>ya-y</td>
</tr>
</tbody>
</table>
nouns, verb stems, adverbials and postpositions; and adverbials derived from nouns, verb stems and adverbs. In this section I will first describe some verb derivational affixes (§3.2.1), followed by nominalizing affixes added to verb stems to form nouns (§3.2.2), and then some adverbial derivations from noun and verb stems (§3.2.3).

3.2.1 Verb derivational affixes

In most Carib languages, only one verb derivation category occurs as a prefix: the detransitivizer, which is added to a transitive stem and carries the meanings of 'reflexive' or 'reciprocal', or simply 'intransitive' which is often best translated as a passive in languages like English. There are many categories which occur as suffixes. Table 2.9 shows the more frequently occurring ones grouped into three main types: valency changers (causative of an intransitive verb and causative of a transitive verb), aspect markers (ingressive, terminative and cessative), and word class changers, which derive verbs from nouns; one set of suffixes forms intransitive verbs, and the other two sets form transitive verbs with meanings like: benefactive or productive for one set, and malefactive or reversative for the other set. Many of the forms in the eight languages are clearly cognates.

Ergative MA (Abbott 1991) has one construction that is quite different from the other languages: the formation of causatives of both transitive and intransitive stems is primarily by means of a separate verb, emapuhti 'cause' - the subject of which is the causing agent - and a preceding subordinate clause in which the causee is the subject:

(11) imakuhi pe amiiri es-enuninka emapuhti-ya badness DENMLZR 2:PRO DETR-think cause-3-ERG
'She will cause you to think in a bad way.'

MA has a few intransitive verb stems which take a CAUS suffix to form a transitive stem; the suffixes are cognates of some of the forms in table 2.9: -nipi, -pa. There is another suffix which occurs with transitive verbs, -melpa, with the specific meaning of 'order (to do something)'. For other derivations, MA has affixes which are cognates of forms in table 2.9: prefix -eh-, ex-, at- 'Detrzr'; and suffixes -pialti 'Ingressive', -areitiha 'Terminative', -ma 'Intr.vblr', -qa, -ma, -hti 'Tr.vblr', -ka 'Vblzr:mal'.

In all the languages with CAUS suffixes, where the suffix is added to an INTR to form a TR, the S of the INTR becomes the O of the TR and a new A is introduced; where the suffix is added to TR1 to form TR2 in TR, a new A is introduced, the O remains the same, and the A of TR (the causee) becomes an oblique agentive phrase, marked with the agentive/ergative postposition (see §3.1.1, re table 2.6; and

3.2.2 Nominalizing affixes attached to verb stems to form nouns

Most nominalizers are suffixes (one significant exception, the prefix no, is discussed below), and cognate forms regularly appear in the languages for which data are available. Gildea (1998) presents tentatively reconstructed proto-Carib forms of several nominalizers. Table 2.10 is not an exhaustive set of nominalizers, but shows Gildea's reconstructed proto-forms and the cognate forms which are described in my sources. The functions of the suffixes in table 2.10 are: *-ne 'Nominalizer of A'; *-ri 'Nominalizer of Action/State'; *-tipi or *-tupi 'Nominalizer of Action/State; Past Tense'; *-apo 'Nominalizer of S/O, Past Tense'; *-tupi 'thing, time or place.
<table>
<thead>
<tr>
<th>Lg</th>
<th>Prefix</th>
<th>Valency change</th>
<th>Aspect</th>
<th>N+ SUFF = INTR</th>
<th>Word class change</th>
</tr>
</thead>
<tbody>
<tr>
<td>DETRZR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP</td>
<td>e-</td>
<td>ma</td>
<td>-po</td>
<td>-ta</td>
<td>(h)ka</td>
</tr>
<tr>
<td></td>
<td>os-</td>
<td>-nohpo</td>
<td></td>
<td>-pa</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ot</td>
<td></td>
<td></td>
<td>-to</td>
<td></td>
</tr>
<tr>
<td></td>
<td>at-</td>
<td></td>
<td></td>
<td>-po</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>(w)e-</td>
<td>ka</td>
<td>-po</td>
<td>-ta</td>
<td>-ka</td>
</tr>
<tr>
<td></td>
<td>wos-</td>
<td>-nopi</td>
<td></td>
<td>-ma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>wot-</td>
<td></td>
<td></td>
<td>-ka</td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>e-</td>
<td>ho</td>
<td>-tũ</td>
<td>-ma</td>
<td>-tũ</td>
</tr>
<tr>
<td></td>
<td>öt-</td>
<td></td>
<td></td>
<td>-ka</td>
<td></td>
</tr>
<tr>
<td></td>
<td>t-</td>
<td></td>
<td></td>
<td>-tõ</td>
<td></td>
</tr>
<tr>
<td>HI</td>
<td>e-</td>
<td>ka</td>
<td>-ho</td>
<td>-(n)ta</td>
<td>-(n)ka</td>
</tr>
<tr>
<td></td>
<td>os-</td>
<td>-nsh</td>
<td>-mohso</td>
<td>-ma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ot-</td>
<td>-noh</td>
<td></td>
<td>-ka</td>
<td></td>
</tr>
<tr>
<td></td>
<td>at-</td>
<td></td>
<td></td>
<td>-ha</td>
<td></td>
</tr>
<tr>
<td></td>
<td>as-</td>
<td></td>
<td></td>
<td>-(b)to</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>e?-</td>
<td>ma</td>
<td>-pa</td>
<td>-ta</td>
<td></td>
</tr>
<tr>
<td></td>
<td>et-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ot-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>at-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>e-</td>
<td>-nipo</td>
<td>-po</td>
<td>-ke(pi)</td>
<td>-ta</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-kei</td>
<td>-ma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-pa</td>
<td>-ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-hte</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-nte</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>e-</td>
<td>-ka</td>
<td>-fo</td>
<td>-ta</td>
<td>-ka</td>
</tr>
<tr>
<td></td>
<td>es-</td>
<td>-nef</td>
<td>-fo</td>
<td>-ma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>et-</td>
<td>-nofu</td>
<td>-meki</td>
<td>-fa</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ej-</td>
<td></td>
<td></td>
<td>-ka</td>
<td></td>
</tr>
<tr>
<td></td>
<td>et-f-</td>
<td></td>
<td></td>
<td>-fo</td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>e-</td>
<td>-nēp</td>
<td>-po</td>
<td>-ta</td>
<td>-ka(i)</td>
</tr>
<tr>
<td></td>
<td>ēt-</td>
<td>-nip</td>
<td>-nehpo</td>
<td>-ta</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ēh-</td>
<td>-mēpka</td>
<td>-mēpka</td>
<td>-te</td>
<td>-pai</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.10 Nominalizing suffixes added to verb stems

<table>
<thead>
<tr>
<th>Language</th>
<th>-ne</th>
<th>-ri</th>
<th>-tupu-ri</th>
<th>-sapo</th>
<th>-topo</th>
<th>-pini</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>-ne</td>
<td>-ri</td>
<td>-hpiri</td>
<td>-topo</td>
<td>-pini</td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>-nen</td>
<td>-ri</td>
<td>-xpo</td>
<td>-xpo</td>
<td>-topo</td>
<td>-bin</td>
</tr>
<tr>
<td>DE</td>
<td>-dı́</td>
<td>-hö-hë</td>
<td>-töho</td>
<td>-hüni</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HI</td>
<td>-ne</td>
<td>-(ni)-ri</td>
<td>-thiri</td>
<td>-saho</td>
<td>-t(i)ho</td>
<td>-hini</td>
</tr>
<tr>
<td>KA</td>
<td>-ne</td>
<td>-ri</td>
<td>-ti-pi</td>
<td>-töho</td>
<td>-pin</td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>-nen</td>
<td>-api</td>
<td>-sah</td>
<td>-toh</td>
<td>-pin</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>-pe(n)</td>
<td>-(ni)-0</td>
<td>-bpü-hë</td>
<td>-sa?</td>
<td>-to?</td>
<td>-pi</td>
</tr>
<tr>
<td>TR</td>
<td>-ne</td>
<td>-ri</td>
<td>-hipé</td>
<td>-topo</td>
<td>-hini</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>-ne</td>
<td>-ri</td>
<td>-thiri</td>
<td>-topo</td>
<td>-hini</td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>-ne</td>
<td>-0</td>
<td>-tpü/tpi</td>
<td>-topo</td>
<td>-pin</td>
<td></td>
</tr>
</tbody>
</table>

Associated with the state or action; I add one other, not included by Gildea: *-pini

Nominalized Negation).

HI examples are:

(14a) -pe ‘Nominalizer of A’
  ro-hananir-pe
  I-instruct-A.NMLZR
  ‘one who instructs me’

(14b) -(ni)-ri ‘Nominalizer of Action’
  INTR: a-wanotar-niri
  2-sing-AC.NMLZR
  ‘your singing’
  TR: i-hananih-ri
  3-instruct-AC.NMLZR
  ‘the instructing of him’

(14c) -thiri ‘Nominalizer of Action, Past Tense’
  i-wanotar-thiri korno
  3-sing-AC.NMLZR COLL
  ‘their singing (in the past)’

(14d) -saho ‘Nominalizer of the S/O, Past Tense’
  S: i-manho-saho uro/omoro/moki
  IMPERS-dance-S.NMLZR 1/2/3PR
  ‘I (am) / you (are) / he (is) the one who danced.’
  O: t-ono-saho koso
  IMPERS-eat-O.NMLZR deer
  ‘The deer (was) what was eaten.’

The above examples demonstrate the following characteristics that generally apply in Carib languages. The (a) -pe nominalization applies only to transitive verbs; the corresponding intransitive S nominalizer (‘doer of the action’) is a different construction (see §3.2.3 for the nominalization of the -...-so adverbial stem); thus A and S are treated differently. The same is seen in the (d) -saho and (f) -hini nominalizations – the subject of the verbless clauses refers to the underlying S and O of the verb that has been nominalized, never to A. The pivot of the other three nominalizations (b, c, e) is the referent of the person-marking prefix (which functions as a possessor in these nominal constructions) is the underlying S or O of the nominalized verb, never the A. This highlights the ergatively organized characteristic of nominalizations in Carib languages.

There is one construction that occurs in many Carib languages as a nominalization in which both the A and O are in focus. Gildea (1998) refers to this as ‘the idiosyncratic O Nominalizing Prefix n-’. As that label indicates, the pivot of the construction is a prefix n- which refers to the underlying O. There is, however, also a person-marking prefix (or possessor noun phrase) that precedes it which always refers to the A (but which in the surface nominalized form functions as possessor). It usually co-occurs with one of the Action Nominalizing suffixes, including Past, and in PA it can occur with other suffixes (Gildea 1998). I illustrate again from HI in (15a, b):

(15a) n(i)- ‘O Nominalization’
  ro-n-ari-ho-niri
  1-O.NMLZR-take-CAUS-AC.NMLZR
  ‘the thing I send (lit. cause (someone) to take)’
In CA, the _n_-prefix does not function as a nominalizer, but is dependent on the co-occurrence of a nominalizing suffix. Also in CA, the _n_- can refer to the action as well as to the underlying O (Hoff p.c.)

3.2.3 Some adverbial derivations from noun and verb stems

`Adverb' is the term I use for what I describe for HI in Derbyshire (1985: 10) as: a heterogeneous class comprising what in other languages are often distinguished as adjectives and adverbs, and including also locative and time words, and numerals and quantifiers. They are bound together into one class on the grounds of their common syntactic function, either (i) as complement of the copula in copular clauses, or (ii) as adjunct [i.e., oblique constituent] in any type of clause.

There is no recognizable class of `Adjective' in HI; what seem like adjectives, on a semantic basis, function as adverbs or nouns (nominalized forms of adverbs). The above definition covers most of the Carib languages for which I have seen descriptions, although some linguists use the label `Adjective' for some forms and may have good reasons for making finer distinctions for the languages they have described. Postpositional phrases, at least in HI, have the same syntactic functions as, and certain similar morphological properties to, adverbs.

There is a relatively small set of basic, non-derived adverbs, perhaps as many as 100, with meanings such as `good', `tall', `high', `how', `one', `here', `over there', `far away', `with difficulty', `soon', `many', `all'. The antonyms of many of these are formed by adding the suffix -hira `Neg'. There is the potential for thousands of derived adverbs, since there are at least 15 derivational processes (in HI; and most of them are reported for other Carib languages); some of these processes are highly productive. I will describe a few of them, with WW data (16a-d, 17, 18) taken from Hawkins (1998).

There are several adverbializing processes which take the following form: ti- + N or V stem + Suffix. The prefix _ti_- is invariable (except for phonological conditioning); the stem can be that of any semantically appropriate noun or verb; the form of the suffix differs for each process.

(16a) _ti-...sol-fi_ 'state or process, derived from V'

  INTR: t-wayih-so n-e-a-si
  ADVZ-die-ADVZ 3S-be-SF-NONPAST
  `He is dying.' (refers to S of the action)

This derived adverb is followed by the copula, and it is always in the third person singular form, for both intransitive and transitive verbs; this gives it something of an impersonal meaning, as indicated in the English translations. In the case of the N stem, however, the copula is inflected for the appropriate person.

(16b) _ti-...fore_ `obligation/fitness/goodness, derived from V or N'

  INTR: ti-to-fore n-e-a-si
  ADVZ-go-ADVZ 3S-be-SF-NONPAST
  `It would be good to go.'

  TR: ti-enta-fore n-e-a-si
  ADVZ-hear-ADVZ 3S-be-SF-NONPAST
  `It ought to be heard.'

  N: ti-phi-fore m-e-a-si
  ADVZ-wife-ADVZ 2S-be-SF-NONPAST
  `You have a good/beautiful wife.'

(This derived adverb is followed by the copula, and it is always in the third person singular form, for both intransitive and transitive verbs; this gives it something of an impersonal meaning, as indicated in the English translations. In the case of the N stem, however, the copula is inflected for the appropriate person.)

(16c) _ti-...rel-yel-`te_ `state of being, derived from possessed N'

  tu-wuJ1u-fe
  ADVZ-laziness-ADVZ
  `lazy'

(16d) _ti-...ke_ `possessing an entity, derived from possessed N'

  tu-hme-ke
  ADVZ-egg-ADVZ
  `having eggs'

In the forms derived from verbs in (16a) and (16b), the adverbializing suffix has as its pivot the S or O constituent of the underlying verb, as in the case of the nominalizing suffixes (see §3.2.2).

The nuclear constituents of a negative clause are an inflected form of the copula preceded by a phrase containing an adverb that has been derived from a verb stem by the addition of the suffix -hira `Neg' (the WW form; all Carib languages known
Desmond C. Derbyshire

to me have cognates of this form. Compare the Nominalized Negation suffix in table 2.10 (WW -\textit{hi}). Possessed noun stems and some postpositions and adverbs are negated with the same suffix. With forms derived from intransitive verbs and possessed noun stems, the impersonal prefix \textit{yi-}\textit{-lo} co-occurs with the suffix; with forms derived from transitive stems a personal prefix co-occurs, and it refers to the O of the verb; in both cases the inflected copula (or other main verb which occasionally occurs) has the prefix which refers to the subject of the action or state. Examples from WW include:

(17) WW: \texttt{-hra-ra ‘negation adverbializer, derived from V or possessed N’}

\begin{tabular}{l}
INTR: yi-wink\-i-ra w\-a-a-si \\
IMPERS-sleep-NEG 1-be-SF-NONPAST \hfill ‘I will not go to sleep.’ \hfill \\
TR: k\-en-i-hra n\-a-a-si \\
1 + 2-see-NEG 3-be-SF-NONPAST \hfill ‘He does not see us.’ \hfill \\
POSSD.N: yi\-phi\-hra ki\-wIf\-fe-si \\
IMPERS-wife-NEG 1S-go-SF-NONPAST \hfill ‘I’m going without my wife.’ \hfill \\
POSSD.N: e\-ewu-hra w\-e-a-si \\
IMPERS-eye-NEG 1-be-SF-NONPAST \hfill ‘I’m blind / have poor sight.’ \hfill 
\end{tabular}

Another adverbial derivation common in Carib languages relates to temporal and conditional expressions. There are two adverbializing suffixes in WW which mark the nuclear constituent of such expressions: \texttt{-taw ‘when, while, if’}; and \texttt{-\textit{tihe ‘after, if’}. The resulting adverbs have person-marking prefixes or preceding free form nouns that refer to the S of an intransitive, and to the O of a transitive, verb. WW examples are:

(18a) \texttt{-taw ‘when, while, if’}

\begin{tabular}{l}
INTR: tuuna mok\-ya-taw to-hra t\-a-a-si \\
rain come-SF-if go-NEG 1 + 2S-be-SF-NONPAST \hfill ‘If it rains (if rain comes) we will not go.’ \hfill 
\end{tabular}

(18b) \texttt{-\textit{tihe ‘after, if’}

\begin{tabular}{l}
INTR: a\-feaka-t\textit{tihe ti-hIf\-e-si \\
2S-wake-up-after 1 + 2S-go-SF-NONPAST \hfill ‘After you wake up we will go.’ \hfill 
\end{tabular}

TR: o-nom-t\textit{tihe marya fe w\-a-a-si \\
1-leave-after knife wanting 1S-be-SF-NONPAST \hfill ‘After (you) leave me, I will want a knife.’ \hfill 

3.3 The particle word class

Carib languages abound in particles. Most of these are always postposed to other constituents, never occurring alone or as the initial constituent in a phrase. They are never inflected, nor can they have derivational suffixes added (both of which are possible with other types of postpositions, e.g., the locative postpositions described in §3.1.4). I agree with Hoff’s (1990: 495) statement, in defending his use of ‘particle’ rather than ‘clitic’: ‘Even though the Carib elements show a few clitic-like features, these are not sufficient to deny them word status.’ Some are like clitics in that they are phonologically attached to the preceding constituent and are subject to the same morphophonological processes that apply at morpheme boundaries within words. Others are not affected by those processes. All, however, are more mobile than bound affixes and their general syntactic behaviour justifies their being defined as a distinct word class.

They fall into distinct sub-classes on the basis of their meanings and their syntactic and discourse roles. Linguists have labelled and described them in different ways. In Derbyshire (1985) I distinguished three sub-classes for HI: (1) modifier particles, which function almost like adjectives to modify the immediately preceding constituent; (2) discourse particles, which usually refer to something in the discourse-pragmatic context; and (3) verification (i.e. evidential) particles, which express the speaker’s attitude to what they are saying, including degree of certainty and authority. In any particle sequence in HI, members of each class usually occur in that same order; any verification particle is always in the final position in the phrase. Hoff (1986, 1990) distinguishes two classes for CA: (1) non-modal particles, including both my modifier and my discourse particles, and (2) modal particles, which include my verification particles and a few others.

The following examples are from HI (most other languages have similar categories and sometimes what appear to be cognate forms): (1) modifier: \textit{komo ‘Collective’}; \textit{ymo ‘Augmentative’}; \textit{heno}, which has two meanings: ‘Deceased (referring to a person), and ‘Large group’ (referring mainly to animals and plants; e.g. \textit{torono heno} ‘flock of birds’); (2) discourse: \textit{haxa ‘Contrast’}; \textit{rma ‘Same referent’ or ‘Continuity’}; \textit{harha ‘Return to former state or location’}; and (3) verification: \textit{(ha)\textit{ti} ‘Hearsay’}; \textit{(ha)\textit{ha} ‘Uncertainty’}; and \textit{(ha)\textit{mi} ‘Deduction’} (which the speaker makes from known facts).

3.4 Pronouns

Table 2.11 displays the pronoun sets from nine languages. The first set of third person forms are referential, in the sense that they usually refer to entities that have already been introduced into the discourse. The other third person pronouns are
Table 2.11 Pronouns, singular (noncollective) forms

<table>
<thead>
<tr>
<th>Person</th>
<th>AP</th>
<th>CA</th>
<th>DE</th>
<th>HI</th>
<th>MA</th>
<th>PA</th>
<th>TR</th>
<th>WW</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>iwi</td>
<td>a:wi</td>
<td>üwü</td>
<td>uro</td>
<td>uro</td>
<td>yu</td>
<td>wi</td>
<td>oyi</td>
<td>yu</td>
</tr>
<tr>
<td>1+2</td>
<td>kimoro</td>
<td>khikko</td>
<td>kivi</td>
<td>kivro</td>
<td>uurikko</td>
<td>yuto</td>
<td>ki:me</td>
<td>kiwi</td>
<td>kimmé</td>
</tr>
<tr>
<td>1+3</td>
<td>ina</td>
<td>a:na</td>
<td>jnaa</td>
<td>amma</td>
<td>amna</td>
<td>ana</td>
<td>ainya</td>
<td>amma</td>
<td>emma</td>
</tr>
<tr>
<td>2</td>
<td>omoro</td>
<td>amoro</td>
<td>omóodo</td>
<td>omoro</td>
<td>amirí</td>
<td>amén</td>
<td>emén</td>
<td>amoro</td>
<td>emé</td>
</tr>
</tbody>
</table>

3 REFERENTIAL

<table>
<thead>
<tr>
<th>ANIM</th>
<th>INAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>inoro</td>
<td>iro</td>
</tr>
<tr>
<td>itaw</td>
<td>ijóó</td>
</tr>
<tr>
<td>móró</td>
<td>moro</td>
</tr>
<tr>
<td>noro</td>
<td>iro</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
</tbody>
</table>

3 DEICTIC: PROXIMATE

<table>
<thead>
<tr>
<th>ANIM</th>
<th>INAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
</tbody>
</table>

3 DEICTIC: MEDIAL

<table>
<thead>
<tr>
<th>ANIM</th>
<th>INAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
</tbody>
</table>

3 DEICTIC: DISTAL

<table>
<thead>
<tr>
<th>ANIM</th>
<th>INAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
<tr>
<td>móró</td>
<td>móró</td>
</tr>
</tbody>
</table>


demonstratives and usually have the deictic function of relating the person or entity directly to the extralinguistic context of the utterance (as also do the first and second person pronouns). The third person deictic pronouns have two parameters for animacy (animate and inanimate) and three for spatial scaling (proximate, medial and distal). CA has one form for proximate and distal animate. MA has the same forms for medial and distal, and lacks '3 REFERENTIAL.INAN' (the sources are not specific, but it seems that 3 DEIC forms can be used with referential, as well as deictic, meaning).

4 SYNTAX

In this section I describe some of the more distinctive characteristics of Carib syntax under three main headings: main clause structuring (§4.1); subordinate clause constructions (§4.2); and the ergativity which, to a greater or lesser degree, governs the case marking, person marking, derivational processes and constituent order patterning in Carib languages (§4.3). Much of the syntax is found in the morphology, which is described in §3.

4.1 Main clause structuring

There are six basic clause types in Carib languages: transitive (with A and O arguments); intransitive (S argument); copular with an inflected copula; copular with an auxiliary verb (occurs mostly in the languages with main clause ergativity); verbless copular; and quotative (having a main verb 'to say' and an embedded clause of direct speech that normally precedes the 'say' verb).

There is one dominant feature in the ordering of constituents in main clauses: the fairly rigid OV sequence. This correlates with basic phrase constituent orders: GEN-N and N-Postposition. (For the Noun-Modifier type of phrase, see below on paratactic sequences.) What appears to be an exception to the rigidity of the OV order is found in PA, in which VO occurs at least as often as OV, but as Doris L. Payne (1994) points out, this is due mainly to the fact that the final O is often paratactically adjoined to the clause, rather than being an integral part of the clause syntax (again, see below on paratactic sequences). It has proved more difficult to determine whether the basic position for A and S NPs is pre- or post-verbal, since both orders occur, but noun phrases and free pronouns do not often occur once they have been introduced as topics in a discourse. The verb person markers are usually the only indicators of the A and S (and O) referents. Where NPs do occur, the evidence points to transitive AOV and intransitive SV as being basic in some languages (e.g. CA, DE), and VS and OVA as basic in others (e.g. AP, HI, PA). In the ergative languages (e.g. KU, MA) SV and OVA are the dominant orders, although in the transitive clause AOV also occurs. The evidence for these basic patterns is not primarily frequency of occurrence, but the fact that alternate orders can be explained as being pragmatically marked (Derbyshire 1986).

The above applies primarily to the basic configuration in clauses which have only third person referents (A, S, O). Hoff (1995, for the Carib language) and Gildea (1995b, for Carib languages in general) provide insightful studies that show a non-configurational pattern for clauses with first and second person referents, and for certain constructions which, according to Gildea's diachronic hypothesis, are historically more innovative.

There is considerable flexibility in the ordering of oblique constituents (adverbial and postpositional phrases). They usually occur after the verb or any post-verb A/S argument, but can be fronted to clause-initial position for emphasis, contrast, or continuity when referring back to the previous clause(s). There is a
strong tendency not to have more than one constituent preceding the main predication (V or OV).

Paratactic constructions are a feature of Carib sentences and clauses. A sentence may consist of two or more juxtaposed clauses usually in a coordinating function, but without any conjunctions or particles (such as 'and', 'but', 'or', etc.). Only the final clause in such a sequence has the terminal intonation pattern. Within a clause, constituents other than the verb (i.e. A, S and O noun phrases, and adverbial and postpositional phrases) can consist of two or more juxtaposed phrases. The first can usually be considered the head phrase, and the phrases that follow may have one of several functions: coordinating, modifying, adding a greater degree of specification, or as an afterthought clarification. Each phrase in the sequence is usually separated by a brief pause. These sequences can be left-dislocated for emphasis or as a topicalization device. More often they are right-dislocated, following the main predication. A single noun or phrase can also be left- or right-dislocated, as is often the case with the PA post-verbal O constituent mentioned above. There can also be discontinuous sequences in which the head phrase precedes, and the other phrases follow, the verb. For fuller discussion and exemplification, see Derbyshire (1985: 129-35) on HI, and Doris L. Payne (1994) on PA.

4.2 Subordinate clause constructions

In some Carib languages (e.g. CA, HI, WW) there are, strictly speaking, no subordinate clauses. The subordinate constructions take the form of phrases: nominal phrases (§4.2.1), adverbial phrases (§4.2.2) and postpositional phrases (§4.2.3). These phrases have as their nuclear elements nominalized or adverbialized forms derived from verbs. The more strongly ergative languages have a few constructions of that kind, but there are also regular subordinate clauses, including some with finite verb forms. As I pointed out in §3.2.2 and §3.2.3, most of the nominalized forms and some of the adverbialized forms take person-marking prefixes which refer to the S and O of the underlying verb, thus indicating the ergative character of the constructions, and in the derived structures these prefixes express a possessor relationship rather than S or O (although this is sometimes difficult to translate into English). Free form (pro)noun possessors can substitute for third person prefixes. The A of the underlying verb, when it occurs as either a noun phrase or person-marking prefix, is followed by a postposition which has cognates in most Carib languages with the meanings 'to, by, ERG' (HI o-wya in (23c) and romuruwya in (24b)).

4.2.1 Nominalizations

The various types of nominalization and the derivational suffixes (and, in one case, prefix) are described in §3.2.2. Here I give a few examples from WW and MA, of the way these function as A, S and O, and also as modifiers of other noun phrases (corresponding to relative clauses in languages like English) in main clauses. The other functions of nominalized constructions occur in postpositional phrases (§4.2.3). Examples (19a-c) are WW and (20a-b) are MA.

(19a) S of an intransitive clause
miya o-to-Tfow marari y-ama-je komo
away 3-go-COLL + IMM.PAST field GEN-cui-A.NMLZR COLL
'The field cutters went away.'

(19b) O of a transitive clause
a-mok-ri w-enta
2-come-AC.NMLZR 1-hear + IMM.PAST
'I heard you/your coming.'

(19c) modifier of another noun, the whole noun phrase being the A of a transitive clause
p-esk-e-si tak yawaka i-yo-hta-faµu
3S3O-bite-SF-NONPAST now axe IMPERS-edge-VBLZR-S/O,NMLZR.PAST
'The axe that has been sharpened now bites (cuts) things.'

(20a) O of a transitive clause
aw-ennaµho-µpi-kon epuhti-µpi-i-ya
2-return-PAST-COLL know-PAST-3-ERG
'He knew you all returned.'

The derived nominal has a finite verb form but is syntactically nonfinite (see Gildea 1998, chapter 2, for discussion).

(20b) modifier (relative clause with finite tense suffix); post-verbal, right-dislocated, modifying its pre-verbal head (miriri):
miriri erahma-hpi-i-ya ti-rui ni-kupi-hpi
that see-PAST-3-ERG 3REPL-brother O.NMLZR.do-PAST
'See that, what his brother had done.'

4.2.2 Adverbializations

As noted in §3.2.3, some of the adverbs derived from verb stems have person-marking prefixes (or, like derived nominals, a preceding possessor noun phrase) and others do not. The latter usually have the impersonal prefix, which in some languages has the same form as the third person prefix but does not function as such. WW examples are given in §3.2.3 of the more common adverbial derivations. Examples (21a–b) are also WW, showing another frequently occurring derivation, formed by the suffix -sol-ji,
with the meaning 'purpose or goal of a verb of motion'. This derivation is not to be confused with the -sol-§ derivation that co-occurs with the adverbializing ti- prefix (§3.2.3); the two sets of suffixes have the same allomorphy in all the Carib languages for which they have been described (Gildea p.c.), and both result in adverbial functions, but the 'purpose of motion' derivation has person-marking prefixes on transitive verbs and the impersonal prefix on intransitive verbs, while the other derivation has the invariable prefix ti- 'Advz'. Examples from WW are:

WW: -sol-§ 'purpose/goal of motion'
(21a) INTR: yi-win-so ki-wtf-e-si
IMPERS-sleep-ADVZ:PURPOSE.OF.MOTION 1-go-SF-NONPAST
'I am going (away) to sleep.'

(2b) TR: aw-akronoma-fi k-mok-ya-si amyne
2-help-ADVZ:PURPOSE.OF.MOTION 1-come-SF-NONPAST later
'Later I will come to help you.'

MA has the same construction, with the suffixes -il-se, and it is nonfinite. Examples include:

(22a) INTR: suhminan-se toh
play-ADVZ:PURPOSE.OF.MOTION 3:COLL
epahka-piti-hpi poro pona
go out-ITERATIVE-PAST outside to
'They went outside to play.'

(2b) TR: tiaron-kon witi-hpi moroh
other-COLI go-PAST fish
yaphi-se
catch-ADVZ:PURPOSE.OF.MOTION
'Others went to catch fish.'

4.2.3 Postpositions

In postpositional phrases the object of the postposition can be any type of noun or noun phrase, including nominalized constructions. These embedded subordinate constructions are common in Carib languages. Two of the more frequently occurring constructions of this kind in HI will now be described.

In HI, the primary means of conveying the concepts of 'desire', 'want', 'like' and 'love' is with the postposition fe, which has the basic meaning 'desiderative'. It co-occurs with the copula, which carries the inflections for tense and person of the subject. The following examples show that fe: (1) can be inflected with the O person marker (23a); (2) can be preceded by a simple noun and suffixed with the negation adverbializer -hira (23b); and (3) can be preceded by a more complex construction such as a derived nominal (23c).

HI occurrences of fe 'Desid' are:

(23a) o-fe w-ehf-aqa
2-DESID 1S-be-NONPAST
'I love you.'

(23b) a-kanawa-thiri fe-hra w-ehf-aqa
2-canoe-ROSSID.PAST DESID-NEG 1S-be-NONPAST
'I don’t want your old canoe.'

(23c) ro-min yaka ro-katfho y-ok-niri fe
1-house to 1-thing GEN-bring-NMLZR DESID
w-ehf-aqa o-wya
1S-be-NONPAST 2-by
'I want you to bring my things to my house.' (lit. 'I want the bringing of my things to my house by you.')

In (23c) the noun phrase which is embedded as the object of the postposition fe has as its nuclear constituent a possessed noun that is derived from a transitive verb, -oki- 'bring'. The A of that verb is indicated in the sentence-final postpositional phrase omwa by the prefix o- '2'. The root -wya has several usages, including 'to' when referring to recipients or addressees. Here it refers to the A of the nominalized verb, and thus functions as an agentive or ergative marker. In HI, as in many other Carib languages, ergativity operates only in subordinate constructions.

The second embedded construction of this kind is one which occurs with the postposition ke, which has the meanings 'because of', and 'by means of'. HI 'ke' is not inflected for person, nor does it occur with the suffix -hira 'Neg', but it has the same embedding potential as fe, and an embedded derived nominal functions ergatively.

(24a) i-to-hra w-ahko, thenehra tuna
IMPERS-go-NEG 1-be + IMM.PAST much water
y-omok-niri ke
GEN-COME-AC.NMLZR because
'I didn’t go because it was raining heavily.' (lit. 'I didn’t go because of the coming of much rain.')

(24b) k-erwehothe hqko o-wo-niri ke
1-be.happy + NONPAST peccary GEN-SHOOT-AC.NMLZR because
ro-muru wya
1-son by
'I am happy because my son killed a peccary.' (lit. 'I am happy because of the shooting of a peccary by my son.')
I have taken the view that the languages with dominant ergativity represent the earlier stage of Carib morphosyntax and that the other languages, with varying degrees of mixed ergative-absolutive patterning, have developed from that earlier more 'pure' ergative stage (Derbyshire 1991, 1994). This was consistent with a view of constituent order change in the Carib family that I had proposed earlier (Derbyshire 1981), based on a study of constituent order in three languages (CA, HI and MA). In the 1991 paper I compared the ergativity-related characteristics of MA, PA and HI, arguing that the facts of PA, as described in T. E. Payne (1990) and by Gildea (in earlier work culminating in his 1992 dissertation), did not support their reanalysis hypotheses for that language and that the relevant PA constructions could be more plausibly construed as nonfinite subordinate constructions (i.e., not reanalysed as finite clauses). A reference grammar of PA is in preparation (T. E. Payne and D. L. Payne, ms.), and this may help to determine which of the competing views is more convincing for that language.

My view of the direction of change in the Carib family has been reinforced by a more general factor: the rampant ergativity that is found in so many Amazonian language families (Arawá, Carib, Jé, Pano, Tacana, Tupí-Guaraní, Yanomami). This suggests a long history of ergativity in the area (Derbyshire 1987). Preliminary comparisons of some of the linguistic features of Carib and Tupí-Guaraní suggest the possibility of a distant genetic relationship (see Goeje 1909 – cited in Rodrigues 1996, and Derbyshire 1994 for morphosyntactic similarities). Historical and comparative studies are well advanced for Tupí-Guaraní, and Jensen (1998) has made a strong case that the present mixed ergative-accusative systems in the languages of that family can be traced back to more dominant ergativity in proto-Tupí-Guaraní.

A more definitive assessment of the Carib–Tupí relationship (whether genetic or areal), and of diachronic change in all areas of Carib morphosyntax, is dependent on extensive comparative studies aimed at producing reliable phonological and lexical reconstructions and a more reliable internal classification of the Carib family. Such studies can only be undertaken when fuller descriptions become available of more Carib languages.

BIBLIOGRAPHY


Derbyshire, D. C. 1958. 'Kai'chúyana (Questtionalário Padrão para a Pesquisa nas Linguas Indígenas Brasileiras)', Brazil: SIL.


1995a. 'Word order changes in the Cariban family', Ms.


1990. 'Classification of Kariban languages'. Ms.


1996. 'A case of grammatical affinity among Tupi, Karíb, and Macro-Jê'. Ms.


1996. 'Phonological sketch of Waimirí-Atroari (Cariban)', Ms.

1995. 'OVSu versus VSuO in Panare (Cariban): do syntax and discourse match?'. Ms.


1979. Hixkaryalí, Brazil: SIL.


1979. Hixkaryalí, Brazil: SIL.


1979. Hixkaryalí, Brazil: SIL.


The Arawak language family

ALEXANDRA Y. AIKHENVALD

1 INTRODUCTION: THE ARAWAK LANGUAGES AND THEIR SPEAKERS

The Arawak language family contains the largest number of languages in South America. Geographically, it spans 4 countries of Central America – Belize, Honduras, Guatemala, Nicaragua – and 8 of South America – Bolivia, Guyana, French Guiana, Suriname, Venezuela, Colombia, Peru, Brazil (and formerly Argentina and Paraguay).

There are about 40 living Arawak languages. Numbers given in the literature vary from 154 (Loukotka 1968) to 89 (Noble 1965). As is often the case in South America, a single name may be used for what are in fact two or more distinct languages (a single language can also be referred to with several different names). A striking example is the name ‘Baniwa’ which etymologically comes from a Tupi-Guarani term for bitter manioc. It is applied to two languages, Baniwa of Icana (also known as Kurripako) and Baniwa of Guainia, which are about as different as English and Russian. (These peoples do not call themselves Baniwa.) For extinct languages there is often not enough data to make an informed decision. Table 3.1 presents a cautious assessment of what we believe to be distinct languages. Their approximate geographical locations are shown on map 2.

Arawak languages played an interesting role in conquest. The first native American peoples encountered by Columbus in the Bahamas, Hispaniola and Puerto Rico were Arawak-speaking Taino – their language became extinct within the first hundred years of the white invasion (Rouse 1992). Spanish – and thus many other languages – contains a number of loans from Arawak languages, including

1 I am grateful to all my teachers of Arawak languages – Cândido, Graciliano, Olivia, Jovino and José Brito (Tariana), Humberto Bultazar and Pedro Angelo Tomas (Warekena), the late Candelário da Silva (Bare), Afonso, Albino and João Fontes, Celestino da Silva and Cecilia and Laureano da Silva, Januário Paiva and the late Marcílio Rodrigues (Baniwa). Also to David Payne, José Alvarez and Peter van Bamer.

2 The name ‘Warekena’ is also misleadingly used with reference to several distinct Arawak peoples; see discussion in Aikhenvald (1998).
Table 3.1 Arawak languages

<table>
<thead>
<tr>
<th>SOUTH AND SOUTH-WESTERN ARAWAK</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 South Arawak</strong></td>
<td></td>
</tr>
<tr>
<td>1a Te</td>
<td>Te Terêna, East of the Paraguay river, Miranda and Aguidauana rivers in Mato Grosso do Sul, Município do Acai near Bauru, São Paulo; c. 9,800 - G, T, W, D</td>
</tr>
<tr>
<td>1b Ki</td>
<td>Kinikinao, Miranda, Mato Grosso do Sul, Brazil; extinct in the early to the mid twentieth century – W</td>
</tr>
<tr>
<td>1c La</td>
<td>Guané/Layana, Yacare and Galvân rivers in Paraguay, Miranda, Mato Grosso do Sul in Brazil – W</td>
</tr>
<tr>
<td>1d Cha</td>
<td>Chané/Zoomêche, Iiyuro river, Salta Province, Argentina – W</td>
</tr>
<tr>
<td>1e Bau</td>
<td>Bauré, Blanco river, Beni province, Bolivia; c. 5,000 – G, W</td>
</tr>
<tr>
<td>1f Ig</td>
<td>Moxo, or Ignaciano, Mamoré river, Plains of Moxo, Beni province, Bolivia; c. 5,000 – G, D</td>
</tr>
<tr>
<td>1g Tri</td>
<td>Moxo: Trinitario, Mamoré river, Beni province, Bolivia; c. 5,000 – GR</td>
</tr>
<tr>
<td>1h Pai</td>
<td>Palaconce, Sources of Paraguá river, Santa Cruz province, Bolivia – W</td>
</tr>
<tr>
<td>1i Pau</td>
<td>Pauna, Baures river, Santa Cruz Province, Bolivia – W</td>
</tr>
<tr>
<td>1j Apl</td>
<td>Apolitá, Apolobamba, Bolivia; not clear whether it is one language or two – W</td>
</tr>
<tr>
<td>1k Sa</td>
<td>Salumú (Enawené-nawé), Headwaters of Juruena, Mato Grosso, Brazil; c. 154 – W</td>
</tr>
<tr>
<td><strong>2 Pareci-Xingu</strong></td>
<td></td>
</tr>
<tr>
<td>2a Wa</td>
<td>Waurá, Batovi river, Park Xingu, Brazil; c. 130 – GR, W, D</td>
</tr>
<tr>
<td>2b Me</td>
<td>Mehinaku, Me, Dialect of Waurá, between Batovi and Curisevo, Park Xingu, Brazil; c. 95 – GR, W</td>
</tr>
<tr>
<td>2e Ya</td>
<td>Yawalapiti, Curisevo river, Park Xingu, Brazil; c. 135 – P (M), W</td>
</tr>
<tr>
<td>2d Cu</td>
<td>Kustenáu, Batovi and Jatobá rivers, Park Xingu, Brazil; extinct in the twentieth century – W</td>
</tr>
<tr>
<td><strong>Pareci-Saraveca</strong></td>
<td></td>
</tr>
<tr>
<td>2e Pa</td>
<td>Pareci (Haliti), Juba, Guapure, Verde, Papagáio, Burtir and Juruena rivers, Mato Grosso, Brazil; c. 600 (dialects include Waimare, Caximã, Kozarini or Pareci-Cabixí) – G, W, D</td>
</tr>
<tr>
<td>2f Sr</td>
<td>Saraveca, Verde and Paragua rivers, Santa Cruz Province, Bolivia – W</td>
</tr>
</tbody>
</table>
### Table 3.1 (cont.)

<table>
<thead>
<tr>
<th>Region</th>
<th>Language(s)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3 South-Western Arawak</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Piro-Apuriná</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a Pi</td>
<td>Piro (Maniteneri, Maxineri)</td>
<td>Acre, Brazil: c. 1.000; Peruvian Montaña: c. 2.000 – G, W D, GR</td>
</tr>
<tr>
<td>3b Cho</td>
<td>Cho</td>
<td>Iaco and Chandless rivers, Acre, Brazil, a dialect of Piro – W</td>
</tr>
<tr>
<td>3c Ap</td>
<td>Apurina/Ipurina, Cangiti</td>
<td>Along the tributaries of Purús river, southern parts of Amazonas and north of Acre, Brazil; c. 2.000 – GR, W</td>
</tr>
<tr>
<td>3d In</td>
<td>†Hiapari</td>
<td>Madre de Dios, Peru, close to extinction – W, GR</td>
</tr>
<tr>
<td>3e MaPi</td>
<td>†Mashko-Piro</td>
<td>Madre de Dios, Peru, a dialect of Hiapari (David L. Payne 1991: 362) – W</td>
</tr>
<tr>
<td><strong>4 Campa (Ca)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a Asha</td>
<td>Ashaninca</td>
<td>Department of Cuzco, Apurímac/Ene, Tambo, Lower Peru, Peru; c. 15,000–18,000 – G, D, W</td>
</tr>
<tr>
<td>4b Ashe</td>
<td>Asheninca</td>
<td>Department of Cuzco, Pichis, Apurímac, Upper Peru, Ucayali headwaters, Peru; c. 18,000–25,000 – P, M, GR, D, W</td>
</tr>
<tr>
<td>4c Cq</td>
<td>†Caquinte</td>
<td>Department of Cuzco, Peru; c. 200–300 – P, M</td>
</tr>
<tr>
<td>4d Ma</td>
<td>Machiguenga</td>
<td>Cuzco, Andean foothills; c. 7,000–12,000 – GR, W</td>
</tr>
<tr>
<td>4e No</td>
<td>Nomatisiguenga</td>
<td>Cuzco department, Peru; c. 2,500–4,000 – GR, W</td>
</tr>
<tr>
<td>4f PC</td>
<td>Pajonal Campa</td>
<td>Gran Pajonal, Ucayali headwaters, Cuzco, Peru; c. 8,000 – P, M, GR, W; Peru, Pichis and Ucayali may be dialects; other Campa languages may be spoken in Acre, Brazil (Rodrigues 1986: 72: 235)</td>
</tr>
<tr>
<td><strong>5 Amuesha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Am</td>
<td>Amuesha</td>
<td>Department of Cuzco, Peru, Loreto; c. 6,000–8,000 – GR, D in prep., W</td>
</tr>
<tr>
<td><strong>6 Chamicuro</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Cha</td>
<td>†Chamicuro</td>
<td>Department of Loreto, Peru; extinct in second half of twentieth century – T, P (M)</td>
</tr>
<tr>
<td><strong>NORTH-ARAWAK</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7a Wp</td>
<td>Wapishana</td>
<td>Branco, Tacutú, Maú, Surumú rivers in Roraima, Brazil; c. 1,500; and adjacent regions of Guyana: c. 9,000; Atalai, on Rupununi and Cuyuwini rivers, Brazil and Guiana – GR, W, D</td>
</tr>
<tr>
<td>7b Mw</td>
<td>†Mawayana/Mapidian/ †Mawakwa</td>
<td>Roraima, Brazil; about a dozen old speakers in Wai Wai villages – W</td>
</tr>
<tr>
<td>8a Pr</td>
<td>Palikur</td>
<td></td>
</tr>
<tr>
<td>8b Mr</td>
<td>†Marawan</td>
<td></td>
</tr>
<tr>
<td>8c Aro</td>
<td>†Arun/Aroá</td>
<td></td>
</tr>
<tr>
<td>9a IC</td>
<td>Island Carib (Ierí)</td>
<td></td>
</tr>
<tr>
<td>9b Ga</td>
<td>Garifuna (Black Carib, Carift)</td>
<td></td>
</tr>
<tr>
<td><strong>Tet-Arawak subgroup of Caribbean</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9c LAr</td>
<td>Lokono/Arawak</td>
<td></td>
</tr>
<tr>
<td>9d Guaj</td>
<td>Guajibo/Wayyu</td>
<td>Suriname; c. 700; Guyana: c. 1,500; French Guiana: c. 130–200; a few in Eastern Venezuela – GR, G, T, W</td>
</tr>
<tr>
<td>9e A6</td>
<td>A6un/Purahuno</td>
<td>Venezuela, Colombia, Guajiro Peninsula; c. 300,000 – T, D, GR, G, PED GR, W</td>
</tr>
<tr>
<td>9f Tai</td>
<td>†Taino</td>
<td>State of Zulia, Venezuela, Lake Maracaibo; almost extinct – GR, G</td>
</tr>
<tr>
<td>9g Cq</td>
<td>†Caciqueto</td>
<td>Bahama Islands, Hispaniola, Puerto Rico, Cuba; Jamaica extinct within 100 years of conquest; possibly several languages (Rouse 1992: 5) – few words</td>
</tr>
<tr>
<td>9h She</td>
<td>†Shebayo</td>
<td>Venezuelan coast; extinct in mid sixteenth century (Oliver 1989: 55) – few words</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trinidad, a Venezuelan coast; extinct before the twentieth century – few words</td>
</tr>
</tbody>
</table>
### Table 3.1 (cont.)

#### North-Amazonian

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10a</td>
<td>Re</td>
<td>Resigaro, Northeast border of Peru, Bora and Ocaína villages; formerly spoken on the Caquetá river; few speakers in 1970s – G, D</td>
</tr>
<tr>
<td>10b</td>
<td>Yu</td>
<td>Yucuna (Guarú), Miriritaparaná, Amazonas, Colombia; c. 600-700 – GR, W</td>
</tr>
<tr>
<td>10c</td>
<td>Acha</td>
<td>Achaguas, Meta, Colombia; c. 200 – G, W (old GR)</td>
</tr>
<tr>
<td>10d</td>
<td>Pia</td>
<td>Piapoco, Vaupés, Colombia; c. 3,000 – G, GR, W</td>
</tr>
<tr>
<td>10e</td>
<td>Cab</td>
<td>Cabyari, Miritaparaná river, Colombia; c. 50 – W</td>
</tr>
<tr>
<td>10f</td>
<td>Mp</td>
<td>Maipure, Vichada territory, Colombia; W, GR</td>
</tr>
</tbody>
</table>

†Cabe is another language formerly spoken in the Vichada territory, Colombia, for which a short word list is available.

#### Upper Rio Negro

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10g</td>
<td>Blq</td>
<td>Baniwa of Içana/Kurripako, Including dialects: Carru, Carátana, Dzawni, Pauli, Pasu (Cacá), Catapolitani, Holódeno, Sisui; Kumandene (Ipek), Ayacucho, etc.; Içana river and its tributaries, border of Brazil, Colombia, Venezuela; c. 3,000-4,000 – GR, D in prep., G in prep., T, W</td>
</tr>
<tr>
<td>10h</td>
<td>Ta</td>
<td>Tariana, Vaupés river, Brazil; c. 100 adults – GR, D in prep., G in prep., T, W</td>
</tr>
<tr>
<td>10i</td>
<td>Guar</td>
<td>Guaragua, tributary of Içana, Brazil; c. 2; Guzmán Blanco, Venezuela; c. 300 – GR, W</td>
</tr>
</tbody>
</table>

#### Orinoco

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10j</td>
<td>Ba</td>
<td>Bare, Casiquiá, Amazonas, Venezuela, upper course of Rio Negro, Brazil; just a few old speakers left (Guiaú: state of Bolívar, Venezuela: W) – G, D in prep., GR, W</td>
</tr>
<tr>
<td>10k</td>
<td>BGua</td>
<td>Baniwa of Guainia, Including a dialect Wurukana of Xié (WX); Xié river, tributary of Içana, Brazil; c. 10; Guainia, Venezuela; c. 200 – G, D in prep., W</td>
</tr>
</tbody>
</table>

#### Yavítero (Baniwa al’ Yavita)

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10l</td>
<td>Yav</td>
<td>Yaví (Baniwa of Yavita), Village of Yavíta, Atuapa river – G, W</td>
</tr>
<tr>
<td>10m</td>
<td>Mnd</td>
<td>Mandawaska, Baria, Pasionimi rivers in Venezuela; possibly extinct – W</td>
</tr>
<tr>
<td>10n</td>
<td>Yb</td>
<td>Yabaana, Maraula, Cauaboris, Brazil; extinct in the second half of the twentieth century – W</td>
</tr>
</tbody>
</table>

#### Middle Rio Negro

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>10o</td>
<td>Ka</td>
<td>Kaifana, Brazil; a few speakers reported in the 1950s – GR, W</td>
</tr>
<tr>
<td>10p</td>
<td>Mn</td>
<td>Manao, Brazil; extinct in the eighteenth century – T, W</td>
</tr>
<tr>
<td>10r</td>
<td>Bhw</td>
<td>Bahwana/Chiriana, Dementi river, Amazonas, Brazil; 1 speaker reported – G, small D</td>
</tr>
</tbody>
</table>

Other extinct languages of the Río Negro and its tributaries are: †Amarizana (Vera and Agua Blancas rivers, Meta, Colombia), †Amuyu (Amazonas, Venezuela), †Caririu (Río Negro and Pastauri), †Maraté (Icá river, Brazil), †Passé (Negro, Içá, Icá rivers, Brazil), †Waimuna (Icá river, Amazonas, Brazil), †Waraicú/Araikú (Jurua river, Amazonas, Brazil) – W, †Wirina (Marari, Rio Branco, Brazil), †Yumuna (Paré and Joami rivers, Amazonas, Brazil) – W |

### Notes:

Only living languages and extinct languages on which materials are available are included here and in map 2. The map covers only the languages spoken in South America (that is, 9a, 9b and 9f are not included). Endangered languages are marked with t; extinct languages are marked with †. Names which probably refer to dialects are given in parentheses. The second column gives the abbreviations by which languages are referred to in the text, e.g. Te for Terena. The following abbreviations are used to indicate the materials available: T – texts, G – grammar, PED GR – pedagogical grammar, M – morphology, P – phonology, GR – brief grammatical description, W – word lists, D – dictionary.
Languages in areas settled by the European invaders soon became extinct. Those on the north coast of South America perished early, before 1700. When the search for gold and rubber spread up the Amazon and its tributary, the Rio Negro, further languages succumbed. Sometimes the Indians retaliated, attacking settlements and missions; but the invaders always returned. Indian rebellions often provoked forced migrations which sometimes ended up creating a new dialect or even a new language. For instance, in 1797 the British authorities removed the rebellious inhabitants of St Vincent (an island in the Lesser Antilles) to Belize on the mainland. Racially, these were a mixture of Indians and black slaves who spoke an Arawak language called Island Carib. This resulted in the creation of a new dialect of Island Carib – known as Central American Island Carib, Cariff, Black Carib or Garifuna – which by the twentieth century had developed into a separate language, now one of the Arawak languages with the largest number of speakers (Taylor 1977b).

The overwhelming majority of Arawak languages are now endangered. Even in the few communities with over 1,000 speakers, a national language (Portuguese or Spanish) or a local lingua franca (Lingua Geral Amazônica, Quechua or Tucano) is gradually gaining ground among younger people. A massive switch to Lingua Geral Amazônica took place around 1900 in the region of the Rio Negro, and resulted in the rapid loss of a number of languages. Numerous dialects of Baniwa of Içana spoken on the Lower Içana in Brazil are almost extinct. Tucano is rapidly replacing Tariana in the Vaupés river basin; Yawalapiti is yielding to Kamaiurá as a lingua franca of the Xingu Park (Mujica 1992).

The few healthy Arawak languages are Guajiro in Venezuela and Colombia (estimates vary from 60,000 to 300,000 speakers) and Garifuna in Central America (from 30,000 to 100,000). The Campa languages (total estimate 40,000–50,000) are one of the largest indigenous groups in Peru. The majority of materials collected between 1600 and 1900 consisted of word lists, phrases and a few paradigms. For languages which became extinct early (e.g. Taino, Caquetio, Shebayo) just a few words survived in the early Spanish chronicles. There are materials on a few Arawak languages spoken in the northern regions of South America (Achaguas, Maipure, Island Carib mixed pidgin, Lokono Arawak) which go back to the mid seventeenth century. Valuable materials – including word lists, and sometimes grammatical notes, short texts and dialogues (though not all equally reliable) – were collected by travellers (for instance, Wallace 1853, Martius 1867, Von den Steinen 1886, Chauffanjon 1889, Koch-Grünberg 1911, 1928, Nimuendajú 1932).

Most of the materials on Arawak languages collected during the second half of the twentieth century are by SIL linguists. Their quality and quantity varies. At present, no Arawak language – with the possible exceptions of Lokono and

Resigaro – has been provided with a comprehensive grammar (phonology, morphology, syntax), dictionary and text collection. (See table 3.1.)

1.1 Comparative studies, genetic classification and subgrouping

Comparative and historical studies of the Arawak family have a long history. The genetic unity of Arawak languages was first recognized by Father Gilij in 1783, three years before Sir William Jones's famous statement about Indo-European. The recognition of the family was based on a comparison of Maipure, from the Orinoco Valley, and Moxo from Bolivia. He named the family Maipure. Later, it was 'renamed' Arawak by Brinton (1891) and Von den Steinen (1886), after one of the most important languages of the family, Arawak (or Lokono), spoken in the Guianas. This name gained wide acceptance during the following decades.

Comparative studies of Arawak languages initiated by Gilij were continued by Von den Steinen (1886) who proposed the first subdivision of the Arawak languages. He distinguished Nu-Arawak and Ta-Arawak divisions (based on the form of the 1sg pronominal prefix – see table 3.5). Further studies were done by Adam (1890), Brinton (1891, 1892) and others.

The limits of the family were established by the early twentieth century. The Arawak affiliation of a few languages earlier considered 'problematic' has been proved within recent years. These are Amuesha (Taylor 1954b), Resigaro (David L. Payne 1985), Iñapari (Valenzuela 1991), and Bahwana (Ramirez 1992).

Though there are no doubts concerning the genetic affiliation of the Arawak languages listed in table 3.1., problems still exist concerning internal genetic relationships within the family and possible genetic relationships with other groups. Even the name of the family has been a subject of controversy. The majority of native South American scholars use the name 'Arawak' ('Arúak') to refer to the group of unquestionably related languages easily recognizable by pronominal prefixes such as mu- or ta-`1sg', pi-`2 sg', relative prefix ka- and negative ma-. A number of scholars, mainly North Americans, prefer to use the term 'Arawakan' to refer to much more doubtful genetic unities of a higher taxonomic order, and reserve the term 'Maipuran', or 'Maipurean' for the group of undoubtedly related languages (see David L. Payne 1991, Kaufman 1990; and the introduction to this volume). Here I follow the South American practice and use the name 'Arawak' for the family of definitely related languages, following Rodrigues (1986).

Reconstruction, internal classification and subgrouping of Arawak languages is still a matter of debate; further detailed work is needed on both the descriptive and the comparative fronts.

As mentioned in the introduction to this volume, the putative studies of 'Arawakan' by Matteson (1972), Noble (1965) and others are deeply flawed.
Unfortunately, these have been adopted as the standard reference for the classification of Arawak languages, especially among anthropologists, archaeologists and geneticists, influencing ideas on a putative proto-home and migration routes for proto-Arawakan – see the criticism in Tovar and De Tovar (1984).

The earliest classifications – De Goeje (1928), Mason (1950) and even Loukotka (1968) – were predominantly based on the geographical distribution of languages. A comparative vocabulary of 67 Arawak languages with 151 cognate sets (all of mixed quality) was compiled by De Goeje (1928) as a chapter of his grammar of Lokono Arawak. Later comparative studies were made by Shafer (1959) (see the criticism in Taylor 1961, David L. Payne 1991) and Valenti (1986). The first truly scientific reconstruction of proto-Arawak phonology – over 200 lexical items and a few grammatical morphemes – was published by David L. Payne (1991). However, his subgrouping of Arawak languages, which is based on lexical retentions, rather than on innovations, remains open to discussion.

Taylor (1977a, b) was the first to put forward scientific arguments in favour of North Arawak as a separate subgroup. His contribution to Arawak comparative phonology and to comparative studies of Caribbean Arawak (Island Carib, Garifuna, Guajiro, Lokono) cannot be underestimated. A low level reconstruction has been suggested recently (proto-Lokono-Guajiro, see Captain 1991; and proto-Xingu, see Seki and Aikhenvald forthcoming).

The main problem of Arawak comparative linguistics and subgrouping is the lack of adequate data for many languages. Geographical expansion and considerable linguistic diversity within the family pose the problem of distinguishing areal from genetic phenomena. This is crucial for morphological reconstruction. In some cases the source of areal diffusion can be established. Quechuan influences are discernible in Amuesha (Wise 1976). Resigaro was strongly influenced by Bora and Ocaina (Witoto family). The grammar of Tariana was drastically restructured under the influence of East Tucano languages during the past 300 or 400 years (Aikhenvald 1996a). In other cases the origin of a substratum is not known. The Arawak languages of Peru show structural similarity to neighbouring non-Arawak languages (Harakmbet, Witoto). Palikur is said to have arisen from the merger of 8 dialect or language groups (Diana Green p.c., 1996); these unknown substratum languages may account for the fact that it is one of the most divergent North Arawak languages, and has lost a number of common Arawak features.

The history of the Arawak language family offers interesting examples of language contacts. The creation of a ‘mixed’ language of Arawak–Carib origin in the Lesser Antilles is one of the most surprising pieces of evidence on language history in pre-conquest times. Speakers of Iñeri, a dialect of the Arawak language now (misleadingly) known as Island Carib, were conquered by Carib speakers. Their ‘mixed’ Arawak–Carib pidgin survived until the seventeenth century (Taylor 1977b, Hoff 1994). In this pidgin, the ‘speech of men’ and the ‘speech of women’ were distinguished in the following way. Women used morphemes and lexemes of Arawak origin, while men used lexical items of Carib origin and grammatical morphemes mostly of Arawak origin. This pidgin coexisted with the Carib used by men and the Iñeri used by women and children. This diglossia gradually died out with the spread of competence in Iñeri/Island Carib among both men and women. However, as a result, Island Carib underwent a strong lexical and grammatical influence from Carib.

A number of grammatical phenomena distinguish North Arawak from non-North Arawak languages (see also Tovar and De Tovar 1984: 120–45). Though there are a few comparative studies of various subgroups (Wise 1986, 1990a, 1991a and b; Derbyshire 1986; Captain 1991), further investigation is needed to decide whether this division is genetic, or is due to different patterns of areal diffusion, and exactly what the subgrouping is. For the time being, we can only be certain of subgroupings of Arawak languages on a very low taxonomic level (e.g. Xinguan languages, South Arawak of Brazil, Pi-Ap). The preliminary subgrouping in table 3.1 is based on this areal-geographical principle.4

The linguistic argument in favour of an Arawak proto-home located between the Rio Negro and the Orinoco river, or on the Upper Amazon (advocated by Lathrap 1970, Oliver 1985) is the higher concentration of structurally divergent languages in this region than in other Arawak-speaking areas. This is highly suggestive and corroborated by a few mythical traditions concerning a northern origin among Arawak-speaking peoples in southern regions (e.g. in Xingu). The origin myths of the Tariana and Baniwa in the north also suggest that they came from the Caribbean coast. The expansion of Lokono-Iñeri from the Northern Amazon to the islands in the Caribbean is estimated to be quite recent (just a few hundred years before the conquest).

2 PHONOLOGY

2.1 Segmental phonology

A core of segmental phonology found in most of the languages is presented in tables 3.2 and 3.3. The likely phoneme system of proto-Arawak comprises these consonants

3 He established 18 phonetic correspondence sets between 36 Arawak languages. This was a start; unfortunately most of these data were of mixed quality and strongly biased towards North Arawak. Of the 36 languages, 13 are, in fact, dialects of Baniwa of Içana/Kurripako, and only 3 languages are not North Arawak.

4 In all subsequent lists of Arawak languages I follow the order given in table 3.1.

5 Note, however, that Oliver's scheme is based on Noble's (1965) and Matteson's (1972) classifications and his own lexi­co-statastic comparisons of highly unreliable data from 30 languages.
A full set of aspirated voiceless stops is found only in some North Arawak languages

(6) A glottal stop is only found in Te, Bau, Ig, Cha, Re; Wa, Mw, Guaj; Yu, Acha, Pia; Ka; Bhw. Glottalization is a word prosody in Wa, Ya and Ta (§2.2). Aspirated voiceless stops may have been innovated independently by different Arawak subgroups (pace David L. Payne 1991). Aspirated voiced labial and dental stops are found only in Ta; they result from vowel fusion across a morphological boundary, e.g. *di-hpo' (3sg.nfem-take) > *dhpo' 'he takes'.

(7) Bilabial fricatives are rare. Voiced labiodental r is found in Pi, Ap, Ca and Ga, and Ka. Bhw has â. Am has a voiced ß. Only Re distinguishes f and r.

(8) Most languages have two sibilant fricatives. Some North Arawak languages have just one: it is r in Ga, LAR, Ya, Pla, and in Mw; Añ, BX; Ta. A voice contrast in sibilant fricatives is found in Re (alveolar z and s) and in Am (alveopalatal ß and s). Bß and Acha have alveopalatal ß (derived from *r before a high vowel). Alveopalatal fricatives are retroflexed in Bß (voiced) and in Am (voiceless). A distinctive innovation of Ya is an alveolar fricative f (from proto-Xingu *s, or from *tj Seki and Aikhellvald forthcoming). Am is unusual in having a voiced velar fricative y.

(9) A glottal fricative h, or a velar fricative x, has restricted occurrence in Ap, Wp, Añ, Bß, Ta, Ba, WX (see §2.3) and may alternate with zero in Bç, Ba and Bhw.

(10) Most South and South-Western Arawak languages (lg, Wa, Ya, Pa, Pi, Ap; Ca; Am, Cha, Re) have two affricates, while North Arawak languages tend to have either alveolar or preaspirated fricatives; both LAR and Ga have a voiceless Cricative. A distinctive innovation of Ya is an alveolar fricative f (from proto-Xingu *s, or from *tj Seki and Aikhellvald forthcoming). Am is unusual in having a voiced velar fricative y.

(11) A typical Arawak system is similar to other languages of Amazonia in having a single liquid phoneme with a flap and/or, more rarely, lateral articulation, as in Bau, Ig, Ya, Ap, Ca, Mw, Wp, Añ, Ya, Acha, Pi, WX, Ba. Bau and Ig have an alveolar vibrant, and Añ and Ga have a trill. Palatalized laterals are found in some Ca dialects and in Ya. There are four alternative systems:

(i) one rhotic, one lateral, as in Te, Wa, Pa, Pi, Am, Cha, Pr, Ga, Ka, Ye;
(ii) two rhotics, no lateral: a lateral flap and a trill in Guaj; dental and velar fricatives in Bß, and simple and preaspirated fricatives in Bß;
(iii) three liquids: flap, tap/trill and lateral in LAR and trill, flap and lateral in Ta;
(iv) neither a rhotic nor a liquid: Re (Allin 1975: 67), where g, d, n are reflexes of proto-Arawak *r, possibly as the result of areal diffusion from Witoto.

(12) Ba, Bß and Ta have preaspirated nasals, while Re has voiceless nasals. These are historically derived from a metathesized sequence, nasal + h, a process similar to the one described in (6) above, e.g. Bß *kuwaihna > *na-ha '1sg-EMPH'.

(13) n is often unstable in syllable-initial and syllable-final positions. In Te it is replaced by

Table 3.2 Composite statement of consonants in Arawak languages

<table>
<thead>
<tr>
<th>labial</th>
<th>dental</th>
<th>alveolar</th>
<th>laminoo-alveo-</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiced stop</td>
<td>(b)</td>
<td>d</td>
<td>k</td>
<td>(?)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>(g)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspirated voiceless stop</td>
<td>(p')</td>
<td>(g')</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless affricates</td>
<td>s</td>
<td>f</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>palatal</td>
<td>ts</td>
<td>Tj</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glottals</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasals</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glides</td>
<td>w</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3 Composite statement of vowels in Arawak languages

<table>
<thead>
<tr>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i, i:</td>
<td>a, u, u:</td>
</tr>
<tr>
<td>middle</td>
<td>e, e:</td>
<td>a, a:</td>
</tr>
<tr>
<td>low</td>
<td>a, a:</td>
<td></td>
</tr>
</tbody>
</table>

The following comments expand on tables 3.2 and 3.3.

(1) North Arawak languages have some variation of places of articulation: only bilabial, apical, velar and glottal are found in LAR. Añ has bilabial, dental, palatal, velar and glottal. Bß has the maximum of seven: bilabial, dental, alveolar, alveopalatal, palatal, velar and glottal. There is an additional set of palatal stops in Re (palatalized alveolar stops). Bß (palatalized dental stops) and Am (palatalized bilabial stop). All stops have a palatalized counterpart in. Perene Ca (J. Payne 1989). Bß distinguishes dental and alveolar stops. Wp may have glottalized stops.

(2) Phonemic voice distinctions in bilabial and dental stops are found in most North Arawak languages and, possibly, in No (Wise 1990); they are absent from most of the other languages. Dieastically, stops are unstable; thus, p > h in Ya, Pa, Acha, Mnd, Mn, and p > *f in Bß; t > *a before a front vowel in Yu, Acha, d > h or * in Am and Ba; d > *a before a back vowel in Re, Pia, Ca, Bß, Ta, Yaw-WX; k > *a before a front vowel in Yu, Re, Yu, Acha, Ca, Bß, Ta, k > *a inWX/ BGua, and also in Ta, Te, Bw, Pl.

(3) The phoneme b is often restricted to onomatopoeia and a few loans (e.g. Pa, Bß, Ta, Ba); or results from a phonological process (WX); b and d are in a complementary distribution in Bß and Acha (b restricted to p in other languages before *n; p corresponds to h before *a). Mw and Yaw have no bilabial oral stops.

(4) A voiced velar stop is found only in Re and in Ya. In Re it developed from the bilabial and from *a, possibly due to the areal influence of Bor and Wito languages.

(5) A glottal stop is only found in Te, Bau, Ig, Cha, Re, Wa, Mw, Guaj; Yu, Acha, Pla, Ka and Bhw. Glottalization is a word prosody in Ya and Ta (§2.2).

(6) A full set of aspirated voiceless stops is found only in some North Arawak languages (IC, Re, Yu, Bß and Ta). Bß has ph and kh; and dh is extremely rare. These languages developed aspirated labial and velar stops as the result of a metathesis with h (which, in turn, comes from *k before a high vowel). An example of the sequence stop-V-h in proto-North-Arawak is *ku:pa:k 'fish' > *ku:pa:ki > Bß, Ta ku:ph. LAR also has aspirated apical and velar stops, both LAR and Ga have a voiceless fricative developed from ph (David L. Payne 1991). Aspiration in stops may have been innovated independently by different Arawak subgroups (pace David L. Payne 1991). Aspirated voiced labial and dental stops are found only in Ta; they result from vowel fusion across a morphological boundary, e.g. *di-hpo' (3sg.nfem-take) > *dhpo' 'he takes'.

(7) Bilabial fricatives are rare. Voiced labiodental r is found in Pi, Ap, Ca and Ga, and Ka. Bhw has â. Am has a voiced ß. Only Re distinguishes f and r.

(8) Most languages have two sibilant fricatives. Some North Arawak languages have just one: it is r in Ga, LAR, Yu, Pla, and in Mw; Añ, BX; Ta. A voice contrast in sibilant fricatives is found in Re (alveolar z and s) and in Am (alveopalatal ß and s). Bß and Acha have alveopalatal ß (derived from *r before a high vowel). Alveopalatal fricatives are retroflexed in Bß (voiced) and in Am (voiceless). A distinctive innovation of Ya is an alveolar fricative f (from proto-Xingu *s, or from *tj Seki and Aikhellvald forthcoming). Am is unusual in having a voiced velar fricative y.

(9) A glottal fricative h, or a velar fricative x, has restricted occurrence in Ap, Wp, Añ, Bß, Ta, Ba, WX (see §2.3) and may alternate with zero in Bç, Ba and Bhw.

(10) Most South and South-Western Arawak languages (lg, Wa, Ya, Pa, Pi, Ap; Ca; Am, Cha, Re) have two affricates, while North Arawak languages tend to have either alveolar or preaspirated fricatives; both LAR and Ga have a voiceless Cricative. A distinctive innovation of Ya is an alveolar fricative f (from proto-Xingu *s, or from *tj Seki and Aikhellvald forthcoming). Am is unusual in having a voiced velar fricative y.

(11) A typical Arawak system is similar to other languages of Amazonia in having a single liquid phoneme with a flap and/or, more rarely, lateral articulation, as in Bau, Ig, Ya, Ap, Ca, Mw, Wp, Añ, Ya, Acha, Pi, WX, Ba. Bau and Ig have an alveolar vibrant, and Añ and Ga have a trill. Palatalized laterals are found in some Ca dialects and in Ya. There are four alternative systems:

(i) one rhotic, one lateral, as in Te, Wa, Pa, Pi, Am, Cha, Pr, Ga, Ka, Ye;
(ii) two rhotics, no lateral: a lateral flap and a trill in Guaj; dental and velar fricatives in Bß, and simple and preaspirated fricatives in Bß;
(iii) three liquids: flap, tap/trill and lateral in LAR and trill, flap and lateral in Ta;
(iv) neither a rhotic nor a liquid: Re (Allin 1975: 67), where g, d, n are reflexes of proto-Arawak *r, possibly as the result of areal diffusion from Witoto.

A few languages (Bau, Ig, Pia) have ß (Baptista and Wallin 1968, David L. Payne 1991) in free alternation with w before -l, -e and -o and in some forms before suffixes. Guaj also has ß in free alternation with w in some cases (José Alvarez p.c.).
and Pi are unusual in yet another way: for dapa (*kakon) and Wp closed syllables arise from vowel elision in rapid speech registers, e.g. Bau in the south, and Pr in the north allow other consonants in the coda position. In Ball and C] can only be a glide (wor).

In most languages syllable structure is (C)V(V)(C), where C can be any consonant, excluding those in parenthesis. These will need to be confirmed by systematic reconstruction. (They are similar to the set proposed by David L. Payne 1991.)

2.2 Syllable structure, suprasegmentals and phonotactics

In most languages syllable structure is (C)V(V)(C), where C can be any consonant, and C can only be a glide (w or y), a nasal, or, rarely, a liquid or h. Only Am and Cha, in the south, and Pr in the north allow other consonants in the coda position. In Bau and Wp closed syllables arise from vowel elision in rapid speech registers, e.g. Bau dêkolon-akon ‘full’ (Baptista and Wallin 1968: 9), Wp, normal speech , slow speech dap ‘house’ (Manoel Santos p.c., 1993). Consonant clusters are found only in Pr, Wa and Pi are unusual in yet another way: ? or h is inserted before a word-initial vowel.

There are a few phonotactic restrictions. The phoneme h often occurs only word-initially or word-finally as a word boundary marker (Aikhenvald 1996b). In Guaj the glottal stop cannot occur at word boundaries. Word-initial liquids are found only in cross-referencing prefixes (table 3.5).

Stress is contrastive in some languages (Bau, Wa, Ya, Ba, WX, Acha), but not in others (Ca, Bhw). Nasalized or long vowels are often stressed. Stress assignment may depend on syllable weight (Ashe: J. Payne 1990; Guaj: Alvarez 1994).

There are at least two tone languages – Te in the south, and Re spoken in the far northeast of Peru (formerly spoken on the Caquetá river in Colombia). Nomatsiguenga also distinguishes two tones (David L. Payne 1997). Te has two tones – falling and level. Re distinguishes high and low tones, similar to the neighbouring Witoto languages. Some North Arawak languages of the Upper Rio Negro and Colombia (Acha, Pin, Bic, Ta) have pitch accent.

Glottalization, aspiration and nasalization are often word prosodies in North Arawak and in the Xinguan languages, while h-insertion and transylaryngeal vowel harmony are used to mark a phonological word followed by a pause (Aikhenvald 1996b).

2.3 Morphophonology

The following morphophonological processes are found in Arawak languages, usually at morpheme boundaries within a grammatical word: vowel harmony; metathesis of stop and h; vowel fusion. Metathesis of stop and h yields preaspirated consonants in Re and in a number of North Arawak languages (Allin 1975: 184; Aikhenvald 1995a, 1996b). Nasalization and vowel harmony are word prosodies in Te and in the closely related Kinikinao, Guaná and Chané (Bendor-Samuel 1966), and are used to mark 1st and 2nd person respectively; in these languages prenasalized stops and fricatives arise as the result of nasalization. Many North Arawak languages have progressive transylaryngeal vowel harmony, e.g. Ba wa-+ -ihisa > wahasa (1pl-look for) ‘we look for’.

Tone in Te is associated with a unit word rather than a syllable; it is assigned to one of the first three syllables in a word (the placement depends on the morphological structure – Eckdahl and Butler 1979). Falling tone entails vowel lengthening (Eckdahl and Butler 1979); Te lacks phonologically long vowels. A preliminary comparison of cognates shows correspondences between falling tone in Te and long vowels in some North Arawak languages; e.g. falling tone: Te hê; Bic, Ta ku:phê ‘fish’ (proto-Arawak *kopikê; David L. Payne 1991); level tone: Te akmo, Bic, Ta, WX (*h)emu ‘tapir’ (proto-Arawak *kema; David L. Payne 1991). The nature and origin of tones in Te requires further investigation.
There is the following interdependence between syllables and morphemes:

 prefixes: almost all monosyllabic
 suffixes: mostly monosyllabic, a few disyllabic
 roots: mostly disyllabic, a few monosyllabic.

A morphological word usually contains at least one root morpheme; monosyllabic verbal roots have an obligatory prefix (or a suffix, to make them bimoraic, as in Asheninca: David L. Payne p.c., 1997). Thus, every grammatical word contains at least two syllables.

3 TYPOLOGICAL PROFILE

3.1 Word structure

All Arawak languages are polysynthetic and predominantly agglutinating with a few elements of fusion. They are mostly head-marking. This accounts for the lack of core cases used for marking grammatical relations in the majority of Arawak languages. Elements of dependent marking found in Ta, in the north, and in Ap, from the Piro-Apurina subgroup, may be due to areal diffusion.

Arawak languages usually have a rich noun class and/or classifier system and a separate system of two or three genders restricted to pronouns and cross-referencing markers. Nominal morphology is fairly similar across the family. Verb morphology varies more and can be very complicated, especially in South Arawak and South-Western Arawak languages, and also in some North Arawak languages (Guaj, Ta).

3.2 Morphological processes

Arawak languages are mostly suffixing, with just a few prefixes. These are: A/S, pronominal prefix to verbs and possessor prefix to nouns; relative-attributive ka- and negative mo- to nouns and verbs. Causative and applicative prefixes on verbs are found in South Arawak, Pareci-Xinguan and Peruvian languages. All types of prefixes are historically less stable than suffixes; there is much more diversity even among closely related languages.

Many Arawak languages have reduplication of stem-initial CV- (Te, Pr), or

8 All nominal stems which contain more than two syllables in the proto-Arawak reconstruction proposed by David L. Payne (1991) can be decomposed into a disyllabic root and a suffix, or are loans.

3.3 Word classes

All Arawak languages have nouns and verbs; most also have an open class of adjectives. The size of the adjective class varies from language to language. Adjectives share numerous properties with verbs; in Guaj they can be considered a subclass of stative verbs. In some North Arawak languages (Ta, Blc) adjectives share some properties with nouns and some with verbs.

9 In one case partial suppletion of a possessed form may go back to proto-Arawak. The common Arawak term for 'house, home' typically displays an alternation pëpan; cf. Terena pëti, possessed pëna 'house'; Ignesiano peti, possessed pena; Baniwa pan-ti, possessed -pana; Taraita pan-iti, possessed -pana; Palikur poti, possessed -piti; Bahwana panti-iti, possessed -pana; Pareci hali, possessed -hama 'house' (David L. Payne 1991 reconstructs *pe and *pana as two different items).
(v) Only a few Arawak languages — IC (Taylor 1956: 31-2), Acha (Meléndez 1989), Bhw (Ramirez 1992: 46) and Pr — have a small set of generic classifiers employed in possessive constructions with alienably possessed items. Bhw has one classifier for game, and one for domestic animals, e.g. *nu-iRa habuRu (1sg-POSS:CL:DOMESTIC ANIMAL) 'my parrot'.

Classifiers show great diversity from one language to another in semantics and form. A classifier in one language often corresponds to a lexeeme, or to a derivational suffix, in another. Classifiers appear to have developed on the level of individual subgroups.

Typical pronominal genders are masculine and feminine. Ig and Pr also have a neuter gender.¹² No genders are distinguished in the plural.¹³ The markers are uniform across languages, and they go back to proto-Arawak third person singular cross-referencing markers (table 3.5).

Feminine is the functionally unmarked gender in the Caribbean Arawak languages Guaj, Añ, LAr and IC/Ga; in the other languages masculine is unmarked. Gender assignment is rather opaque in these languages, as well as in Ca and Pr; elsewhere it is straightforward: feminine gender is used for females, and masculine in all other cases.

Pronominal genders have been lost in Te, Am, Cha, Pa, Wa and Bhw; all these languages do have classifiers.¹⁴

4.3 Number
All Arawak languages distinguish singular and plural on nouns. Typical markers are reflexes of proto-Arawak *(a)ni-* 'animate/human plural', *(a)pe 'inanimate/animate non-human plural'. Pucal is optionally distinguished from plural in WX and BLC. Re is the only Arawak language with a category of dual in independent personal pronouns and in nouns; this must have evolved under the influence of Bora and Ocaina (Allin 1975: 164). There is a residue of dual in Guaj in verbal cross-referencing on about forty verbs (José Alvarez p.c.).

In a few North Arawak languages — Ta, BLC and Re — nouns with an inanimate referent have a collective meaning; a noun classifier in a derivational function has to be added for singular reference, e.g. Ta *deri 'banana (collective)', *deri-pl (banana-CL:LONG) 'a banana'. Only a noun with a classifier can be pluralized, and verb is the most complicated part of the grammar, and the only obligatory constituent of a clause.

18 In Ca, Am and Pr different particles are used to form pronouns of different persons; e.g. Ashe 1sg *-askaln-aro, 3sg *-ir-iro(ri); Am 1sg na; 3pl *-eh. In Pi 1sg and 3sg pronouns do not involve a pronominal prefix.

19 Demonstrative stems are always monosyllabic. In almost every language third person pronouns can be used as proximate demonstratives, and as definite articles with nouns. Most languages have a two-term system, near/far (Wp, Re, Yu, Acha, Pi, BLC, WX) or a three-term one, near/mid-distant/far (Pi, Ca, Ta, LAr, Ba). Am has just one demonstrative pronoun (*ai 'this/that'; Wise 1986: 572), and Guaj and Pr have four: near speaker and hearer, far from speaker and near hearer, far from both, very far from both.

Interrogatives vary even among closely related languages; they can often be used as relative pronouns; in Ca, Am, Ta, BLC they also function as indefinites.

Most languages have just the numbers 'one' (proto-Arawak *pa- and *two' (proto-Arawak *(a)pi and *yama: David L. Payne 1991). It is common to use numeral *pa- 'one' in the function of an indefinite pronoun 'one, someone', 'another'. South Arawak, Ca, Am, Cha and Pr have reflexes of *mapa 'three' (Payne 1991: 421); other languages show different forms (Pa *hannana; Ta, BLC, Acha maia, mada; Guaj ajanin, LAr kabyin; Wp idikinay/day, Mw itukuura; Ba kiakumana (cf. Apolista erikoni)). The proto-language may have had only 'one' and 'two'. Only a few languages (Pr, LAr, Guaj) have underven numbers up to ten.

Adpositions (postpositions and prepositions; see §7.2) are often derived from body parts; there are some underven adpositions. An adposition in one language can correspond to a verbal valency-changing suffix in other languages, e.g. Te *yaa 'with, in, to', *yol 'dative/locative applicative'; Pa -kakoo, Pr *kak 'with', Pi, Ap, Te, Ig, Ta -kaka 'reciprocal; associative'; South Arawak and Pr-*api nilin- 'comitative', Ca -imo 'circumstantial applicative (in the presence of)', North Arawak languages -ina 'with' (cf. §6.3).

6 VERBAL MORPHOLOGY
The verb is the most complicated part of the grammar, and the only obligatory constituent of a clause.

18 Pi has two genders, and the relative-interrogative pronoun kla also has a special form for inanimate gender.
19 There are two exceptions to this: Ap also distinguishes two genders in 3rd person plural; Ta has an optional distinction of two genders in all persons in the plural in independent pronouns.
20 Wa distinguishes two genders in demonstratives only; Pa distinguishes two genders in nominalizations, and Te has an opposition animate/inanimate in adjectives and descriptive verbs.

e.g. *deri-pl (banana-CL:LONG) 'bananas'. This feature is also found in neighbouring languages, Guahibo and Tucano (see chapter 14 below).

5 CLOSED CLASSES: PERSONAL PRONOUNS, DEMONSTRATIVES, INTERROGATIVES, NUMBERS, ADPOSITIONS
Personal pronouns are used to emphasize the subject in topicalization, and in copula clauses. They consist of a cross-referencing prefix plus an emphatic one-syllable particle — see table 3.5, e.g. Pa na-iya, Ba na-ri, WX nu-yu, Ta nu-ha 'I'.'¹⁸

Demonstrative stems are always monosyllabic. In almost every language third person pronouns can be used as proximate demonstratives, and as definite articles with nouns. Most languages have a two-term system, near/far (Wp, Re, Yu, Acha, Pi, BLC, WX) or a three-term one, near/mid-distant/far (Pi, Ca, Ta, LAr, Ba). Am has just one demonstrative pronoun (*ai 'this/that'; Wise 1986: 572), and Guaj and Pr have four: near speaker and hearer, far from speaker and near hearer, far from both, very far from both.

Interrogatives vary even among closely related languages; they can often be used as relative pronouns; in Ca, Am, Ta, BLC they also function as indefinites.

Most languages have just the numbers 'one' (proto-Arawak *pa- and 'two' (proto-Arawak *(a)pi and *yama: David L. Payne 1991). It is common to use numeral *pa- 'one' in the function of an indefinite pronoun 'one, someone', 'another'. South Arawak, Ca, Am, Cha and Pr have reflexes of *mapa 'three' (Payne 1991: 421); other languages show different forms (Pa *hannana; Ta, BLC, Acha maia, mada; Guaj ajanin, LAr kabyin; Wp idikinay/day, Mw itukuura; Ba kiakumana (cf. Apolista erikoni)). The proto-language may have had only 'one' and 'two'. Only a few languages (Pr, LAr, Guaj) have underven numbers up to ten.

Adpositions (postpositions and prepositions; see §7.2) are often derived from body parts; there are some underven adpositions. An adposition in one language can correspond to a verbal valency-changing suffix in other languages, e.g. Te -yaa 'with, in, to', *yol 'dative/locative applicative'; Pa -kakoo, Pr *kak 'with', Pi, Ap, Te, Ig, Ta -kaka 'reciprocal; associative'; South Arawak and Pr-*api nilin- 'comitative', Ca -imo 'circumstantial applicative (in the presence of)', North Arawak languages -ina 'with' (cf. §6.3).
The majority of prefixes and suffixes in table 3.5 are listed in Matteson (1972) and David (6). Suffix (5) A prefix for non-focused 1sgn is unique in that both suffixes and prefixes can be used on a number of factors: grammatical function, aspect, and the discourse prominence of the argument. Prefixes are used for A and S, if the verb is in progressive aspect or

and as the pronounal possessor to nouns; while O (direct object) is marked with the same suffixes as S (subject of intransitive stative verbs). Among these split ergative languages are Te, Ig, Bau, Pt, Ap, Bhw, Ka, LR, Bñ, and WX. A typical Arawak split ergative pattern is illustrated with examples (4)–(7), from Bñ.

Example (4) shows a cross-referencing prefix which refers to A, and a suffix which refers to O.

(4) ri-kapa-ni
3sg.nfem + A-see-3sg.nfem + O
‘He sees him/it.’

In (5) the same prefix is used for S, on a verb of motion; and in (6) it is used to indicate possessor.

(5) ri-emhani
3sg.nfem + S-walk
‘He walks.’

(6) ri-tfimu-ni
3sg.nfem + POSSR-dog-POSSV
‘his dog’

In (7), the suffix refers to the S argument of a stative verb ‘be cold’.

(7) hape-ka-ni
be.cold-DECL-3sg.nfem + S
‘He is cold.’

However, this pattern is either partially lost or reinterpreted in many individual languages. For instance, in Pr the split ergative pattern survived only with interrogative predicates (Aikhenvald and Green 1998: app. 1). Pr is also extremely unusual in that both suffixes and prefixes can be used for all of A, S, and O depending on verbal aspect (Green and Taylor 1972).

In Campa languages, the choice of cross-referencing prefixes or suffixes depends on a number of factors: grammatical function, aspect, and the discourse prominence of the argument. Prefixes are used for A and S, if the verb is in progressive aspect or

<table>
<thead>
<tr>
<th>person</th>
<th>sg</th>
<th>pl</th>
<th>sg</th>
<th>pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>nu- or ta-</td>
<td>-na-</td>
<td>-te</td>
<td>-wa</td>
</tr>
<tr>
<td>2</td>
<td>(p)i-</td>
<td>-pi</td>
<td>-pi</td>
<td>-hi</td>
</tr>
<tr>
<td>3nfem</td>
<td>ri-</td>
<td>-ri-</td>
<td>-i</td>
<td>-ri</td>
</tr>
<tr>
<td>3m</td>
<td>thu-</td>
<td>-thu-</td>
<td>-u</td>
<td>-thu</td>
</tr>
<tr>
<td>'impersonal'</td>
<td>pa-</td>
<td>-pa</td>
<td>-a</td>
<td>-pa</td>
</tr>
<tr>
<td>non-focused A/S</td>
<td>-i</td>
<td>-i-</td>
<td>-a</td>
<td>-ni</td>
</tr>
<tr>
<td>dummy S/O</td>
<td>-n</td>
<td>-n</td>
<td>-ni</td>
<td>-ni</td>
</tr>
</tbody>
</table>

Notes:
(1) First person singular forms n- and - may have coexisted as variants in proto-Arawak. Their distribution is as follows: -t- is found only in the extreme north (LR, Guaj and AB), while all the rest have n-. Exclusive/inclusive oppositions are atypical for the Arawak languages. In Ca (Reed and Payne 1983) and Te, the original 1sg now means both 'I' and 'we (excl)', while the original 1pl form now has an 'we (incl)' sense. In Ba and Ta the 1sg impers are used in the sense of 1st inclusive.

(2) Second and third person plural forms underwent modification across the family. Combinations of a prefix and a suffix are used for 2pl in Cha, and for 3pl in Pi and Ap. 2pl and 3pl fall together in Ba (one) and WX (ni). Pia has lost the reflex of *i(i) and uses the combination of -p- and -v instead (Klump 1995), while Te has lost pi-.

(3) Te, Am, Wa, Pa use a zero prefix for 3sg. Both forms are used in Bñ, and a few other northern languages; their choice depends on the discourse prominence of the A/S argument (Aikhenvald 1995b).

(4) The evidence for the proto-Arawak origin of *pa- 'impersonal' (also used as reflexive and as 'same subject' marker) is as follows: Wa pV-, Pa ha- '3rd possessor cleftial with the subject', Am pP-'3rd possessor' (Wise 1986: 570); Wp pa- (Tracy 1974: 123-4), Mw a-, Bhw f'p- 'same third person subject', Guaj pa- 'reciprocal'; Yu pa-... (chicha) 'reciprocal' (Wise 1990); Bñ, Ta, WX pa-, Ba, Guaj ha- 'impersonal'.

(5) A prefix for non-focused A/S is found only in the North Arawak languages. The distribution is as follows: -t- is found in Bñ and Ta (Aikhenvald 1995a), Pr r- 'indeterminate possessor', and possibly Pia r-21 while -t- is attested in Guaj, Añ, Ba; Pr; and in LAr and IC (Taylor 1977b: Peter Van Baarle p.c.).

(6) Suffix -ni is found in Pa -t-e/ as 'object marker', Wa -n/aphoristic 3rd person object pronoun (Richards 1991: 163); Am -t-em/ surface object marker (Wise 1986: 572). Bñ and Guaj have merged *-ni and -*t-. Bñ masc. al- fem. ar- masc. al- fem. ar- Guaj masc. nT- fem. w- masc. nT- fem. w- Pr has a neuter nT-'definite' as well as a- (see Aikhenvald and Green 1998, for discussion).

21 The majority of prefixes and suffixes in table 3.5 are listed in Matteson (1972) and David L. Payne (1991). Impersonal *pa-/pa- was added by Wise (1990).

22 In Cha, closely related to LAr, have 1sg n- Cha innovated a-t- for 1sg (an influence from 1pl?). In Wp 1sg is -a-. In Terena and closely related languages (Guaná, Chán, Kinikinako), the A/S, 1sg prefix nu- has become a nasalization prosody, which continues until the first stop or fricative; it yields nasals vowels and prenasalized stops.

23 In Pia, -t- 3sg.masc is homophonous with a verbal cross-referencing prefix used with a
if this argument is topically discontinuous and less foregrounded, while suffixes are used for S, and O, and also for S, if it is topically continuous and foregrounded, or is in stative aspect (David L. Payne and J. Payne 1991; also see Dixon 1994: 211).

Languages which have preserved suffixes just for 3rd person continue using prefixes for A/S. Among these, Wa, Pa and Wp use a suffix to mark O, while Acha, Yu and Pia use 3rd person suffixes to cross-reference a new topic in A/S function.

Guaj and Ahi have two kinds of verbal conjugation: suffixal (or 'analytic') and prefixal (or 'synthetic'). Cross-referencing prefixes in an analytic verbal conjugation refer to the subject (A/S) of a verb which is low on the transitivity hierarchy. Cross-referencing prefixes used in a prefixal conjugation with transitive verbs with a referential definite object (Alvarez 1994: 91) refer to the A of a verb high on the transitivity hierarchy, and cross-referencing suffixes refer to the O. In subordinate clauses, prefixes mark A/S, while suffixes mark O/S.

Ya, Re, Ta, Ba have lost their cross-referencing suffixes.25 Re, Ya and Ta have no split ergativity; Ba and WX have elements of a S/O pivot (also see §7.6). Ta has innovated a case system of a nominative-accusative type (see §7.1). Example (8) illustrates object case marking on a pronoun in Ta.

(8) diha di-ka-makha nu-na
the:SUBJECT 3sg.mfem-see-REC.PAST.NON. VISUAL 1sg.OBJECT
'He has seen me.'

Some North Arawak languages (WX, LAr, Acha, Blq, Ta, and Ba) have a small class of intransitive verbs, the S of which is marked with a dative adposition or with an oblique case (Ta). They refer to physical and mental states. Ga (Munroe ms.: 6ff.) has several subtypes of 'oblique' subjects (marked with prepositions): dative, instrumental and benefactive, used with verbs denoting physical or mental states. Transitive verbs of knowing and liking take dative subjects.

6.3 Valency-changing derivations

Most Arawak languages have both valency-increasing and valency-decreasing derivations. Valency-increasing derivations are more complicated in South Arawak, Pareci-Xingu, Piro-Apuríná, Campa and Amuesha than they are in North Arawak.

Arawak languages have morphological causatives. The valency-increasing (causative) markers found throughout the family are (-)l(-j); a(-)ta(-), (-)ml(-)/, (-)ka(-) (Wise 1990). North Arawak languages tend to have just one derivation,24 while South Arawak and Pre-andine usually have more. The possibilities are:

(i) causatives of intransitive or ambitransitive verbs are marked differently from those of transitive verbs (e.g. the distinction of i- and ko-ika-in Te).

(ii) in Am and a few Ca languages a causative prefix and a suffix are used when the causer is directly involved, otherwise only a suffix occurs.

(iii) forceful causation may be distinguished from non-forceful causation, as in Ma; Pi, and Ap (Wise 1986, 1990). Some Ca languages have a special causative 'make do by proxy' (Shaler 1971: 33).

Within the North Arawak branch, Guaj (-ira), Ahi (-yera), IC (-god(a)), LAr (-xyty), and Re (-tu) form causatives on both transitive and intransitive verbs.27 In other languages, morphological causatives are formed on intransitive verbs28 only (Acha -da, Yu, Bhw, Ba, WX -ta, Pia -idalda; Ta, Blq -i(-tu)). Periphrastic causatives are then employed to causativize transitive verbs.

South Arawak, Piro-Apuríná, Campa and Amuesha have very complicated applicative derivations which put an oblique constituent into the O slot. These valency-increasing derivations apply to comitative, dative, benefactive or malefactive, reason, purpose, instrument, allative, elative, prepositional (i.e. in the presence of somebody) and to any 'thematic' argument (i.e. the one with reference to which the action is done).29

The oblique constituent which becomes O is cross-referenced on the verb, as shown in (9), from Asha (Kindberg 1961: 537):

(9) y-owetfik-i-ni-ro
3masc-make-NON.FUT-BEN-3fem
'He made it for her.'

In most Peruvian languages applicatives are typologically unusual. One verb can contain several applicative markers. The applicative suffixes show that a peripheral

25 In Guaj, causatives cannot be formed on stative verbs.
26 In WX, Ba and Blq causatives are also derived from transitive ingestive verbs (e.g. 'eat, drink').
27 The number of morphemes, and their behaviour, differs from one language to another. Thus, 'purpose' applicative requires a reflexive in Aseh; in No and PC purpose + reflexive indicates purposeless action. In Aseh, presentational applicative -mo is also used as a benefactive and comitative marker unlike other Ca languages (Wise 1990: 111). In Am, -mp forms the thematic argument advancement is often used idiosyncratically, to refer to a vehicle in which the action is performed, or to a meteorological condition.
28 The same morpheme is often used for causative and comitative applicative derivations (e.g. Ta -mo, Wise 1990: 98; Ap mi, Wise 1990: 108; No and other Ca -a-ukal-ag, Wise 1986: 593-4). Comitative applicatives and causative meanings can be ambiguous, e.g. Ma:

no-panki-ag-ak-e-ri
1-plur-EPITHETIC-CAS/GCOM/PERF-NON.FUT-3sg.masc
'I ordered him to plant / I planted with him.' (Wise 1990: 95)
According to the explanation in Olza Zubiri and Jusayú (1978: 228), the main meaning of tend to employ the same marker for reflexives and reciprocals (Aikhenvald forthcoming-e.g. 'thematic' suffixes which mark derivations similar to purposive applicatives, suffix 1990: 100-1) and also reflexives and reciprocals.

(12) nu-huku-naku-t-ene makula
1sg-wash-CAUS-PROLONGED bowl
'I am washing (the inside of) a bowl.'

In contrast, most North Arawak languages have fewer applicatives. Guaj has a suffix -hiraa which marks comitatives (Olza Zubiri and Jusayú 1978: 227-9; Wise 1990: 100-1) and also reflexives and reciprocals.30 Lo (Pet 1987: 46) has non-productive 'thematic' suffixes which mark derivations similar to purposive applicatives, e.g. áiba 'go away', áiba-ha 'pursue', ácuaha 'to go for wood'. Most Arawak languages have one agentless passive. North Arawak languages tend to employ the same marker for reflexives and reciprocals (Aikhenvald forthcoming).

30 According to the explanation in Olza Zubiri and Jusayú (1978: 228), the main meaning of this valency-changing derivation is that 'many subjects take part in the action'.

Example (11), from Pi, shows that one verb can contain two causative markers, two applicative derivations – comitative and dative – and an anticipatory passive (Matteson 1965: 81).

(11) r-u-mun-kak-yehitx-i-ko
3maSC-COMIT-come-CAUS-MANDATORY.CAUS-DATIVE-ANTICIPATORY.PASSV
'He is to be commanded to be caused to come concerning something.'

Pa, Wa, Pi and also Guaj have a different strategy for deriving applicatives. They incorporate adverbials and postpositions; the argument of a postposition then becomes an O or a derived S (cf. the discussion on how an adposition in one language can correspond to a verbal valency-changing suffix in another language, in §5). Example (12) illustrates the incorporation of the postposition nuka 'in' in Wa.

(12) no-p-ako-ts-imo-tsi-ro-ri
1sg-give-REFERENCE.TO-EPENTHETIC.IN-PRESENCE.OF-ASPECT
3sg.fem-3sg.masc
Irema Irocarto paño
Irene Richard scarf
'I gave Richard the head scarf in Irene's presence.'

Passive markers typically contain -k, -n and/or -w; the marking is often quite different even between closely related languages.

Guaj is unusual in that it has two passives: -na and -ur; the latter implies a special effort from the agent (Olza Zubiri and Jusayú 1978: 233-9).

Reflexives and reciprocals are typically marked by verbal suffixes. Reflexive pronouns are found in LAr and Ga, and in Mw.31

6.4 Other verbal categories

All Arawak languages have rather complex systems of tense–aspect, mood, modality, directionals and aktionsarts; only a few have evidentials. These are almost always expressed with optional suffixes or enclitics.32 Languages show great variability in the categories and forms used.33

The majority of South Arawak, Pareci-Xingu and Peruvian Arawak languages have a three-fold aspect distinction: completive (completed, perfective or telic action); progressive (action/state in progress; also a durative meaning); and habitual. A few languages have a future marker. Other aspect-like distinctions include momentaneous, prolonged, imminent action, customary, etc. Peruvian Arawak languages (Ca, Am and Cha) as well as Te and Wa distinguish future/realsis and non-future/realsis (Wise 1986: 566-7).34

Many North Arawak languages distinguish just past (or completive) and non-past. Yu, Acha, Pia -mi 'past, completive' and Re -mi 'recent past' (Allin 1975: 342)

31 Pet (1987) reports that reflexive pronouns are ousted reflexive derivations in young people's LAr; on IC, see Munroe (ms.), Taylor (1977b). This may reflect the influence of neighbouring languages – English, Spanish or Creoles.

32 Aspect marking in obligatory in Te (progressive -ti vs a); it is neutralized in negation and under topicalization.

33 Payne (1991: 380-1) suggests a proto-Arawak origin for a number of directional (*-ape 'arriving', *-ape 'leaving'), aspectual (*-pe 'perfective', *-ena 'progressive, gerund'); *-aka 'perfective, habitual') and modal (*-atu 'future, irrealis', *-al 'conditional', *-ka 'interrogative') suffixes. There are a few more recurring suffix shapes, e.g. Wp -nin 'non-present'; B1g, Ta -nina 'remote past'; Acha -mina 'dubitative' and Bhw -mita 'desiderative'; Pr -ki 'verificational' and Ba -Fki 'verificational, emphatic'. However, more comparative and descriptive work remains to be done to establish the nature of these correspondences.

34 In Caquinte (Swift 1988: 209), non-future reflexive is a portmanteau morpheme (-o), but future reflexive is not (-e-Npa 'future-reflexive').
are cognate with Ta, Blc, WX, Ba -mi 'past, used, pejorative'. Ta is unusual in having a complicated tense-aspect-evidentiality system possibly due to areal diffusion from Tucano (see chapter 14 below).

There is typically just one imperative which involves a zero-marked verbal stem. A few languages have politeness distinctions in imperatives (Pi -si and Acha -ju 'polite imperative', Pr -nu 'intimate request', -ka 'permissive'; Wp -na 'strong imperative'). The rich system of imperative clitics in Ta may again be the result of areal diffusion. Marking of interrogative mood is restricted to Pi, Ca, Guaj and Ta.

A declarative mood marker -ka is a characteristic feature of North Arawak languages (and is also present in Pi).

Ig, Bau, Pa, Wa and the Peruvian languages have rich systems of modalities (e.g. desiderative, reason, frustrating, optative, pretense, purpose, assertive, dubitative – see Ott and Ott 1983; Wise 1986: 598). North Arawak languages, with the exception of Guaj, Pr and Ta, have fewer distinctions; but they often have optative, necessitative, potential, dubitative, conditional and frustrative.

Evidentials are rare, with just a few languages having a marker for 'reported', or 'hearsay' (e.g. Te, Ig, Wa, Pa, Pi; Acha, Re, Pia, Blc, LAr). Bhw (Ramirez 1992: 64–5) distinguishes -bi 'reported, quotational' and -hi 'inferred'. Ta has a rich system of evidentials which emerged as the result of areal diffusion from Tucano. South Arawak and Peruvian languages have a number of verbal affixes which indicate the direction of movement of the subject (see J. Payne 1982 on their temporal and discourse functions, e.g. Te -pol-pol-p, Pi, Ca -ap, Am -alp, Ig -pa 'point of time reference', Bau -piko 'towards', Wa -pe 'towards'; Ashe -an, Acha -ahant 'from'; Am -am, n 'leaving', Bau -na 'action to one place then to another', Wa -ami 'going out'.

These languages also have a great number of aspect-type miscellaneous suffixes which refer either to the manner in which the action is performed (e.g. Te -pl-o 'another time', No -si 'intensive', Pi -m 'distributive', -lева 'characteristic action', Asha -ina 'interrupted', -na 'frequently back and forth', Pajonal -ri 'reversative') or to time (e.g. Ca -man 'early in the morning', -mínk 'late afternoon'). In Ca and Am, verbal suffixes mark number of participants; there is one marker for AS and one for S/O.

Directionals are rarer in North Arawak languages (e.g. Wp ma-... -kan 'do back and forth', Pr -elga 'do while moving', -re 'in', -wa 'far away', -paj 'near by', Re -lah 'towards'). They also have fewer aksiomsart suffixes. Among these are Wp -k 'suddenly', -aan 'interrupted', -dan 'aimlessly', Guaj -pu 'frequentative', -rai 'distributive', -pūna 'do in passing', Añ -cha 'diminutive', -VV 'augmentative', Pia -t'ata 'more or less', Ba -nia 'inchoative'. Ta recently developed a rich system of directionals and other verbal markers from compounded verbs.

6.5 Noun incorporation

Noun incorporation of direct objects and locative constituents is found in Pi and Ca. In Guaj, any possessed noun can be incorporated; the possessor becomes direct object. Example (14) shows incorporation of an O constituent in Pi.

(14) o-kost'fe-kafre-ta
3sg.masc.pick-up-arrest-EPENTHETIC
'He picked up his arrow.' (Matteson 1965: 40)

Pr, from the North Arawak branch, has a very unusual, albeit limited, pattern of body-part incorporation. While in the overwhelming majority of languages in the Amazon incorporated body parts are placed preverbally (e.g. Mundurukú (Tupi), Yanomami, Tupi-Guarani, Pano, Nadëb (Makú)), they occur post-verbally in Pr.

These incorporated body parts have the function of the O of a transitive verb, or the S of a stative verb. If a verb contains an incorporated body part in the O slot (note that body parts in Palikur are obligatorily possessed), the possessor is raised to be direct object. This is a well-known strategy in incorporating languages (type II in Mithun 1984). The possessor may be cross-referenced on the verb with an object suffix or it may be expressed with a full NP, as in (15) (see Aikhenvald and Green 1998).

(15) a-daha ni hakiks-ota bakimmi-ayh
1neut-for-PASS rub-EYE child-PL
'in order to rub the eyes of the children' (lit. eye-rub the children)

6.6 Relative and negative markers

Besides the cross-referencing prefixes, the most stable prefixes in Arawak languages are ka- 'relative, attributive' and its negative counterpart ma- (Matteson 1972: 165; Taylor 1977b). Ka- derives possessive predicates and denominal verbs, e.g. Wp o-ka-dako-in-(a)-nt (3sg.fem-ATTRIB-mouth-REFL-(EPENTHETIC)-REAL.MOOD) 'she shouted' (lit. 'she mouthed': Tracy 1974: 125), or marks the predicate of relative clauses (Blc, Ta).

Prefix ma- is usually the negative counterpart of ka-, e.g. Pa ka-yanitso-ki (ATTRIB-wife-TRANSITIVIZER) 'find a wife for someone', ma-yanitso-ki (NEG-wife-TRANSITIVIZER) 'take away a wife from someone' (Rowan and Burgess 1979: 103); ka-tai 'have fruit' (of a tree), ma-tai 'be without fruit' (Richards 1988: 200); Pi ka-yhi (ATTRIB-tooth) 'having teeth', ma-yhi (NEG.ATTRIB-tooth) 'toothless' (Matteson 1965: 119). It is widespread throughout the family as a general verbal and nominal negator.
Other means of marking negation differ even among closely related languages. The majority of languages use suffixes or clitics; there is typically a separate marker for prohibition (negative imperative). Person/gender/number oppositions are frequently neutralized in negative and/or prohibitive constructions (Pr, Ta, BIc); tense/aspect oppositions are neutralized in negative constructions in Te and Ba.

Different imperative meanings are neutralized in prohibitive constructions in Acha.

## 7 Syntax

### 7.1 Grammatical relations

As we saw in §6.2, most Arawak languages employ cross-referencing on the verb to mark grammatical relations. Only Ta and Ap have core cases linked to the discourse properties of nouns. Suffix -ne in Ta (homophones with instrumental -ne) goes on a focused subject. A clitic -nu of marks topical non-subject. There is also an obligatory marker for pronominal non-A/S. Ap (Facundes ms-b) has an absolutive (SO) case morpheme -li, used to mark a fully affected constituent. There is typically just one core case morpheme in an NP, usually on the last word.

All Arawak languages have peripheral case suffixes (locative, allative, ablative, comitative, instrumental, benefactive). The same morpheme often means both ‘to’ and ‘from’. Languages often have locative/directional -ki-k/ku, locative/ablative -nV, benefactive or instrumental -(m)inV (cf. David L. Payne 1991: 380; Wise 1990). Two peripheral case morphemes often combine in one word, as in Ba -wa ‘perative’ + -uku ‘locative’ -(I)wak ‘along the location’, e.g. nisa-ni-waku (1sg + canoe-nomposs -perative + Loc) ‘along the inside of my canoe’.

### 7.2 Noun phrase structure

In attributive NPs, demonstratives and numerals usually precede the head noun and adjectives follow it; modifiers agree with the head noun in noun class or gender, and in number (if referring to humans or animates). Third person pronouns are often used as definite articles. In possessive NPs, the possessed noun is always the head. The order Possessor-Possessed is found more often than Possessed-Possessor; however, some North Arawak languages have both orders. Inalienably possessed nouns often function as adpositions, the adpositional argument being cross-referenced with A/S/Possessor prefixes.

In many North Arawak languages (Pr, BIc, WX, Ta) the use of adpositions as prepositions or as postpositions depends on the discourse status of the head noun. Postpositions are used when the head noun is not individualized; otherwise prepositions are employed. Example (16), from Pr (Aikhenvald and Green 1998), shows that a postposition is used when the head noun is not individualized, and the adpositional phrase refers to a habitual activity. In (16), the ‘field’ is not individualized. The sentence describes the habitual activity of a woman. Note that -madka is here a postpositional enclitic which forms one phonological word with was.

(16) eg ka-annipwi-yo was-madka
  3fem ATTRIB-work-DUR.fem field-CL:IN.FLAT
  ‘She worked in the field.’

Prepositions are used if the head noun is individualized. Person, number and gender of the head noun are obligatorily cross-referenced on a preposition. This is illustrated with (17). In this example was ‘field’ has nothing to do with any habitual activities associated with a field; the example comes from a story of a Palikur man who was part of a pacification team carefully crossing the Arara Indians’ field while trying to make contact with them.

In Pr, cross-referencing is always obligatory with prepositions whether the head noun is present or absent; but when the same item is used as a postposition, cross-referencing is omitted. In contrast, in Ta and BIc, postpositions take the cross-referencing prefix -’non-focused A/ S’ (Aikhenvald 1995b).

### 7.3 Complex predicates and serial verbs

Auxiliary verbs are a characteristic feature of IC/Ga and LAr. In LAr the auxiliary verb -a takes the cross-referencing and other markers, and the main verb takes suffix -(m) ‘subordinator’ (Pet 1987: 76, 120ff.). This construction is used (a) with a proposed manner adverbial when the adverbial conveys new information – then the auxiliary precedes the main verb, as in (18); (b) with a negated verb, and then the auxiliary follows the main verb, as in (19).
Auxiliary verbs in Te have a very different function. An auxiliary verb construction with \( k\dot{a}e \) is used as a valency-reducing mechanism with an additional meaning of 'provisional state' (e.g. 'be wrapped'), while its causativized counterpart, \( k\dot{i}xo \), is used to mark the causative of a resultative.

Serial verb constructions are an areal property of the languages of the Upper Rio Negro and Colombia. Ba, WX, Blc, Ta, Acha and Pia have productive serialization with motion, posture and modal verbs (see Klumpp 1985: 150; Reinoso 1994, for Pia; Meléndez 1994; Wilson 1992: 32–3; 156–7, for Acha). They express aspectual meanings. Most languages allow no more than two verbs in a serial verb construction. Ta has unusually complicated verb serialization (Aikhenvald forthcoming-c).

Outside this group, Wp has incipient verb serialization (restricted to modal verbs). Motion verbs can be serialized in Pa (Rowan and Burgess 1979: 66–7). And there is limited verb compounding or serialization in Pi (examples in Matteson 1965: 83) and Ca (David L. Payne p.c.).

### 7.4 Constituent order

Wa, Ba and WX have a split ergative pattern in constituent order used to distinguish core arguments: AV0, S, V, VS (see §6.2). In most other languages constituent order has discourse functions. Fronting of focused constituents is found throughout the family.

The South Arawak languages Te, Bau, Ig, Ca and Am, and also Re, LAr, IC/Ga, Guaj and Afi of the North Arawak branch tend to have a verb-initial constituent order. Most languages of the Upper Rio Negro are verb-medial or verb-final; Ta tends to be verb-final as the result of areal diffusion.

Pa, Ya and Ap tend to be verb-medial, while Pi is mostly verb-final.

### 7.5 Non-verbal clauses

In most languages any non-verb can be used as a predicate of existential, locative, equational or attributive clauses; the subject receives \( S_0 \) cross-referencing. Verbalizers and nominalizers are obligatory in just a few languages (e.g. Wp, Te).

Example (20) is from Ig (Ott and Ott 1983: 19) and (21) from Bhw (Ramirez 1992: 52).

(20) \( \text{achichu-havi-ri-pa} \)

tomorrow-1pl.O/S0-COMPL-MOMENTANEous

'Tomorrow we will be (here).'

(21) \( \text{waituranawi-na} \)

man-1sg.O/S0

'I am a man.'

Only very few North Arawak languages show a verb meaning 'have' (e.g. Acha –wâari (Wilson 1992: 120); WX -deka, Ta -de, e.g. \( \text{nu-de-ka t} \) \( \text{f} \) \( \text{f} \) \( \text{inu} \) (1sg-have-DECL dog) 'I have a dog'). There are several strategies for possessive clauses.

(i) Attributive-possessive construction: \( k\dot{a} \)- 'attributive' followed by noun or an adjective is found throughout the family, as in (22), from Te (Ekdahl and Butler 1965: 83) and Ca (David L. Payne p.c.).

(22) \( \text{co-xé'exa-ne} \)

ATTRIB-child-ALREADY

'He/she has children already.' (lit. 'he/she is childed')

(cf. Colloquial English 'He is moneyed')

As we have noted in §6.6, the negative counterpart of \( k\dot{a} \)- is \( m\dot{a} \).

(ii) A construction 'to/for + Possessor-Possessed' is used in a number of North Arawak languages. (23) is from Blc:

(23) \( \text{nu-fiu tfi} \) \( \text{nu} \)

1sg-to dog

'I have a dog.' (lit. 'to me a dog')

(iii) Other North Arawak languages employ the possessed form of a noun in the predicate slot. Example (24) is from Bhw (Ramirez 1992: 52).

(24) \( \text{panera nu-panera-ni} \)

pan Isg-pan-POSSv

'The pan is mine.'

### 7.6 Complex sentences: relative clauses, complement clauses, coordination, pivot and switch-reference

Subordinating verb suffixes are the most common strategy for marking non-main clauses. All the North Arawak languages share a subordinating suffix \( -kal-ko \).

Relative clauses are most frequently posthead, with very few restrictions on the
function of the common argument in either clause. The strategies for marking relative clauses are: (a) relative verb forms; (b) nominalized verbs; (c) relative pronouns; (d) combination of several strategies.

(a) Relative verb forms are found in Te, Ap, Ashe and the majority of North Arawak languages. The predicate of the relative clause is marked with a gender- and number-sensitive suffix (often the same as the 3sg O/S, cross-referencing suffix, in table 3.5). Te uses progressive -ti in this function. Relative verb forms are often used if the common argument is the subject of the relative clause; they agree with the common argument in common gender and number. In (25), from Pia (Klumpp and Burquest 1983: 394), the relative clause is enclosed in square brackets.

(25) áiba astfeli [yà-anè-eri i-pàchí-a-ca] other man 3sg.masc-come-REL:MASC 3sg.masc-visit-ASPECT yà-anà-a i-wènda-Ø amàca 3sg.masc-come-ASPECT 3sg.masc-sell-ASPECT hammock 'The other man who arrived to visit came to sell a hammock.'

In Ashe the cross-referencing prefix on the relativized predicate is deleted if the subject of the relative clause is coreferential with that of the main clause (Wise 1986: 618). Ta (Aikhenvald in prep.), Yu (Schauer and Schauer 1978: 24–5) and LAr (Pet 1987) use different relativizing suffixes on the verb depending on the function of the common argument. LAr has two sets: one (-tha ‘feminine’ and -thi ‘masculine’) is used if the common argument is subject of the relative clause, and the other (-sta) if it is a direct or an indirect object.

(b) In South Arawak languages and in Pa and Pi, nominalized verbs are used to mark the predicate of a relative clause if the common argument is an oblique constituent, otherwise relative forms are used, e.g. Pi niyàwoka (underlying form: -niya-ya-woka 3sg.nfem-eat-LOCATIVE.ADVANCEMENT-SUFF.PLACE) ‘the place where he eats’ (Matteson 1965: 83).

(c) Some languages outside the North Arawak subgroup have relative pronouns (Wa: Derbyshire 1986: 545; Am and Ca: Wise 1986; Pi and No use demonstratives as relative pronouns. Some North Arawak languages (Ta, Ba) use interrogative pronouns as relativizers.

(d) A combination of these strategies is found in some North Arawak languages. For instance, in LAr, posthead relative clauses may have a relative pronoun, while prehead relative clauses cannot have one (Pet 1987: 168). Guaj (Alvarez 1994: 147) marks relative clauses with a demonstrative pronoun ( masc. csi, fem. tìi, pl. na) before the head which is followed by the relative form of the verb ( masc. -kañ, fem. -kañi, pl. -kañan).

A nominalized verb is widely used throughout the family to mark the predicate of a subordinate clause.
constructions to mark A/S as a contrastive topic in Ashe, WX, Bt, Ta, Pia (Klumpp and Burquest 1983: 397).

Several techniques can be used simultaneously. Example (27), from Bt, illustrates the use of cross-referencing prefix i- 'indefinite person' for the fronted A/S, and the relative verb form to make it contrastive:

(27)
\[
\text{h}jia-pia \ i-de-3i-hnua
\]
\[
\text{he-PERF INDEF AS\_take-REL-1sg}
\]

'It was him who took me away.'

**BIBLIOGRAPHY**


Green, H. and Green, D. 1972. 'Surface structure of Palikur grammar'. Brasilia: SIL.

Greene, H. and Green, D. 1972. 'Surface structure of Palikur grammar'. Brasilia: SIL.


Green, H. and Green, D. 1972. 'Surface structure of Palikur grammar'. Brasilia: SIL. Ms.


Green, H. and Green, D. 1972. 'Surface structure of Palikur grammar'. Brasilia: SIL. Ms.


Seki, I. and Aikhenvald, A. Y. forthcoming. 'Proto-Arawak of Xingu'.


1956. 'Island Carib II: word-classes, affixes, nouns, and verbs', IIAJ, 22.1–44.


1974. 'An introduction to Wapishana verb morphology', IIAJ, 40.120–6.


Valenzuela, P. 1991. 'Comprobación del lugar de la lengua Iñapari dentro de la rama pre-
INTRODUCTION

The Tupí family is one of the largest linguistic groups in South America. Nine of its branches are entirely in the Amazon basin while the tenth, Tupí-Guaraní, has most of its languages in Amazonia. This chapter presents an overview of the family, while chapter 5 provides greater detail on the best-known and best-described branch, Tupí-Guaraní. Early comparative work treated the Tupí-Guaraní languages as 'pure' members of the family, and others as 'impure' (Nimuendajú 1948: 214; Loukotka 1950: 26). Later work treated Tupí-Guaraní as just one branch of a larger Tupí family (see Rodrigues 1986 and further references therein). Work is proceeding on a full reconstruction of proto-Tupí (dealing with phonology, morphology and lexicon) and the systematic changes through which modern languages evolved (Rodrigues 1995, Rodrigues and Dietrich 1997); this is required in order to provide a proper proof of genetic relationship.

Table 4.1 lists the languages of each branch (except for Tupí-Guaraní, which is reserved for the next chapter), together with locations and approximate numbers of speakers. II, Aweti, and IV, Mawé, have some similarities with X, Tupí-Guaraní, and further work may be able to show them to constitute an intermediate subgrouping. Map 3 shows the approximate locations of the Tupí languages, excluding Tupí-Guaraní and those languages which are now extinct (marked by 7 in table 4.1). It will be seen that several of the branches now have a single living member. Only a little information has been gathered on Puruborá, which is only remembered as a second language by a few people. Xipáya is remembered by just two old women.

The Mundurukú are one of the best-known tribes in the recorded history of Brazil. During the second half of the eighteenth century they waged fierce warfare against European settlers in the Tapajós region, taking as trophies the heads of their victims. In 1795 (a decade after that other warlike group, the Mura, had surrendered — see § 1.3 of chapter 13) the Mundurukú made peace with the invaders (Hemming 1987: 19–24).
Five of the Tupí branches – Arikém, Mondé, Puruborá, Ramaráma and Tupari – are spoken close together in or near the state of Rondônia. (They have been badly affected by the clearing of forests during recent decades.) This concentration suggests Rondônia as a likely location for proto-Tupí, the putative ancestor language (Rodrigues 1958, 1964).

Languages of the Tupí-Guaraní branch were documented from the sixteenth century. However, languages from other branches only began to be recorded in the nineteenth century. Word lists on Mundurukú were published by von Martius (1867) and Tocantins (1877); on Juruna, Manitsawá and Awetí by von den Steinen (1886, 1894); and on Mawé and Juruna by Coudreau (1897a, b). In the first part of the twentieth century Nimendajú (1924, 1925, 1929a, b, 1930, 1932a, b) contributed good data on Xipáya, Juruna, Kuruáya, Mawé and Arikém.

From the 1950s more systematic work has been attempted, firstly by missionaries with some linguistic training and then by students from Brazilian universities. Some
of the main publications, from which information has been taken for this chapter are:


II Awetí – Emmerich and Monserrat (1972); Monserrat (1976).


2 PHONOLOGY

2.1 Vowels

The most extensive system of oral vowels is in Káro of the Ramaráma branch, with seven members; there are also four nasal vowels – see table 4.2. The Awetí, Mawé and Tupi-Guarani branches have six oral vowels (omitting a from the Káro system) and six nasal vowels. The remaining branches have just five oral vowels (i, e, i, a, o) and a corresponding set of five nasal vowels. Contrastive length (for both oral and nasal vowels) has been reported for the Arikém and Mondé branches, and may also apply in some other branches.

Rodrigues and Dietrich (1997: 268) suggested that proto-Tupí had a six-vowel system: i, e, i, a, u, o. However, correspondence sets involving u and o show an almost complementary distribution (suggesting an original five-vowel system, with [u] and [o] as allophones of one phoneme). Thus o occurs either following a labial consonant (as in *po>Tupinambá (Tb) po ‘hand’), or as the result of a change of proto-Tupí *e before a labialized consonant (*ek>Tb ok ‘house’, *ep>Tb oβ ‘leaf’), whereas u appears in other phonological environments (*k’o>Tb ku‘to eat’, *kop>Tb kuβ ‘to be many’). There are, however, some cases in which u and o contrast (*fok’o>Tb ku’su ‘to chew’, *fjø>Tb so ‘to go’), providing evidence in favour of a six-term proto-system (reconstructing *fjø for ‘to go’, but *fjuk’u for ‘to chew’). Further comparative research is needed.

In the languages of the Arikém branch there has been a cyclic shift of vowels (Rodrigues 1986, Storto and Baldi 1994 and Storto p.c.). Comparing cognates in languages from other branches we find many examples of each of the shifts e>a, a>o, and o>i; and some examples of the shifts i>e, u>i and i>i. It appears that i remained unchanged. Thus, a six-vowel system developed into one with five vowels, but with five of the vowels shifting their values:

<table>
<thead>
<tr>
<th>Oral</th>
<th>Nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front</td>
<td>Central</td>
</tr>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

Source: Gabas (1989)

Table 4.2 Vowels in Káro

Examples include:

- 'bark/shell' Gavião sabee proto-Tupari *ape?
- 'egg' proto-Tupari-Guarani *pe
- 'feather/wing' proto-Tupari *upi + ?a

Further comparative research is needed.

In other branches

<table>
<thead>
<tr>
<th>Karitiána</th>
</tr>
</thead>
<tbody>
<tr>
<td>opa</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>proto-Tupari-Guarani</th>
</tr>
</thead>
<tbody>
<tr>
<td>sipi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>proto-Tupari</th>
</tr>
</thead>
<tbody>
<tr>
<td>papi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>proto-Tupari-Guarani</th>
</tr>
</thead>
<tbody>
<tr>
<td>papi</td>
</tr>
</tbody>
</table>
Aryon D. Rodrigues

'heavy' proto-Tuparí *potsi piti
'jaguar' Káro, Tuparí ameko ombaki
'house' proto-Tupí-Guaraní *kib gep
'moon' Gavião kati oti
'tree' Káro ?ip ?ep
proto-Tuparí *kip
proto-Tupí-Guaraní *7ip

(Note, however, that there are many unexplained forms. For example, 'tree' is liip in Gavião and ip in Mundurukú, with vowel i and not i as in the other three branches just quoted.) Systematic comparative work across all branches of the family will be needed to fully validate and contextualize the proposed Arikém vowel shifts.

There is evidence of vowel shifts o>a and i>t in the Juruna, Mondé and Ramaríma branches but further work is required to investigate how systematic these are.

2.2 Consonants

The number of consonants reported for individual languages ranges from ten to nineteen. One of the largest systems is that in Gavião, from the Mondé branch, shown in table 4.3.

Table 4.4 summarizes the consonant systems of one language from each branch (also repeating Gavião for comparative convenience), excepting Puruborá for which there is no certain data and Tupí-Guaraní, which is covered in chapter 5. (Note that some of these inventories should be regarded as tentative. We have found inconsistencies between different accounts of the phonemes in a given language.) The first row gives those phonemes found in all languages included in the table. Later rows show the additional phonemes found in each specific language. Suruí also has a voiceless lateral interdental fricative, and a voiceless palatal semi-vowel.

Every language in the table has a labial phoneme that may be a semi-vowel or a fricative (or have both realizations). The phoneme represented as /y/ often has a range of realizations. In both Karitiána and Makurap it can be a palatal semi-vowel /j/, nasal /n/ or affricate /dʒ/. (Storto 1997b gives arguments that in Karitiána what we have shown here as y is most appropriately treated as a nasal phoneme.) For Gavião, Moore (1984: 225) states that the phonemic status of /y/ is uncertain since he has recorded no minimal pairs distinguishing it from /s/, /dʒ/ or /n/. Note that in Aweti /s/ is reported to be a retroflex fricative (Emmerich and Monserrat 1972: 9).

Nasals often have a range of realizations, especially in languages with no phone-

<table>
<thead>
<tr>
<th>Table 4.3 Consonants in Gavião</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilabial</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>voiceless stop</td>
</tr>
<tr>
<td>voiced stop</td>
</tr>
<tr>
<td>voiceless affricate</td>
</tr>
<tr>
<td>voiced affricate</td>
</tr>
<tr>
<td>voice fricative</td>
</tr>
<tr>
<td>nasal</td>
</tr>
<tr>
<td>flap</td>
</tr>
<tr>
<td>lateral</td>
</tr>
<tr>
<td>semi-vowel</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Table 4.4 Consonant systems across languages of the Tupí family</th>
</tr>
</thead>
<tbody>
<tr>
<td>in all languages</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>I Karitiána</td>
</tr>
<tr>
<td>II Aweti</td>
</tr>
<tr>
<td>III Juruna</td>
</tr>
<tr>
<td>IV Mawe</td>
</tr>
<tr>
<td>V Gavião</td>
</tr>
<tr>
<td>VI Suruí</td>
</tr>
<tr>
<td>VII Mundurukú</td>
</tr>
<tr>
<td>VIII Káro</td>
</tr>
<tr>
<td>IX Makurap</td>
</tr>
</tbody>
</table>

mic voice stops. Storto (1994) reports that in Karitiána /itä/ has a number of allophonic realizations: (i) as [bmb] between two oral vowels; (ii) as [b] at the beginning of a word, followed by an oral vowel; (iii) as [bm] when following an oral vowel but not followed by another oral vowel (i.e. at the end of a word or followed by a nasal vowel); (iv) as [mb] following a nasal vowel and followed by an oral vowel; and (v) as [m] when between two nasal vowels, or at the beginning or end of a word next to a nasal vowel. The other nasals, /n/ and /ŋ/ behave in essentially the same way. Similar kinds of allophony are reported for Mundurukú and Tuparí and for the Tupí-Guaraní branch (see chapter 5).

For Káro, Gabas (1989) reports that /l/ patterns with /b/ and /g/ (i.e. effectively filling the empty /d/ slot). For instance, there are parallel phonological rules: k→g, p→b and t→r.
2.3 Tone

Although there is no distinctive tone in languages of the Tupí-Guaraní branch, nor in Awetí or Mawé (and probably not in Puruborá) there are two tones in the Monde branch, and in the Juruna language, and there is contrastive pitch accent in Káro. In Karitiana and in languages of the Tuparí branch there is salient pitch accent, but it is predictable from stress which is itself predictable from other phonological and morphological factors. Munduruku has been reported to have four tones (Crofts 1985: xv) but low tone is said to be associated with glottal consonants and high tone is said to carry a message 'many/most', suggesting that there may in fact be just two phonologically contrastive tones.

3 Grammatical Overview

Tupi languages are head-marking and mildly agglutinative. They typically have one pronominal prefix to the verb indicating a core argument and may also have a prefix that marks change of valency. Nouns bear prefixes that mark inalienable and sometimes also alienable possession. Some languages are reported to have a distinct class of adjectives while in others adjectival concepts are coded through a subset of verbs. In Surií adjectives are similar to verbs but differ from them in that only verbs (not adjectives) may take a stative suffix (Van der Meer p.c.).

Moore (1984: 164) reports that in Gavião there are about sixteen verbs that would be translated into English through prepositions, with meanings such as 'be for' (in a benefactive sense), 'be below' and 'be in'. In contrast, Munduruku has a set of postpositions which may follow an NP, e.g. (Crofts 1973: 50; note that raised numbers indicate tones):

(1) \[ \text{má?wit pat lA maygara o wi-n} \]

'man handsome snake kill-ASPECT'

'A handsome man killed the snake.'

Generally, an adjective cannot make up an NP by itself, but it can have as its 'head' just a pronominal prefix. Compare (2) with (3):

(3) \[ \text{a?-pát, máygára wi-n} \]

3sg-handsome snake kill-ASPECT

'A handsome one killed the snake.'

4 Nouns

It appears that a noun in core function within a clause is not marked for case. In some languages peripheral syntactic functions are marked by postpositions, as in (1) from Munduruku, and in others by case suffixes. Tupari has inessive, ablative and allative/instrumental suffixes, e.g.:

(4) \[ \text{te-karo-na wa(p)-m hi-gS-fall-ASPECT hammock-ALL} \]

'He has fallen into the hammock.'

Tupari also has a suffix that marks a constituent which is the focus of that part of the discourse, as in:

(5) \[ \text{te-sito-t o-k-a} \]

3COREF-foot-FOC 3 + O-eat-ASPECT

'He eats his own foot.'

The structure of the noun in Munduruku involves an optional suffix and a possessive prefix. The suffixes are 'diminutive', 'augmentative', 'plural', 'entire', 'specified' and 'dead' (obligatorily used when referring to dead people, rather like the late in English). Thus:
As in other Tupí languages, inalienably possessed nouns (body parts and kin) take one set of pronominal prefixes (e.g. *o'3-ba4 'my-arm') while other nouns may optionally take one of a slightly different set of prefixes marking alienable possession (e.g. *we2-kor be2 'my-canoe') (Crofts 1985: 87).

Most Tupí languages have alternative 3rd person prefixes to a noun according to whether or not the referent is co-referential with some other argument in the clause (as in (5) from Tupari). Compare (7), where the pronominal prefix indicates co-referentiality, with (8), where it doesn't, both sentences being from Káro (Gabas 1994: 137):

(7) na?toA to-wirapo 3sgCOREF-food eat-ASPECT
'The tapir ate its own food.'

(8) na?toA a?wirapo 3sgNON,COREF-food eat-ASPECT
'The tapir is eating its [something else's] food.'

Tupí languages appear to lack grammatical gender but in at least Káro and Mundurukú there are classifiers, which take part in extensive agreement systems. Gonçalves (1987) lists over 120 classifiers for Mundurukú. These function as suffixes to the noun they describe or to a modifying numeral or demonstrative and as prefixes to a verb. For example *ba4 'arm' is also a classifier for deal objects such as a banana. This classifier occurs on each of the words in:

(9) [jep]'lep4-pa6 a'ko1-ba4'o 3sg-su2-ba4'-3o1 two-CL banana-CL 3sgA-3sgO-CL-eat
'He ate two bananas.'

A set of about 10 classifiers is reported for Káro (Gabas 1996). For instance, *kap is used for clusters of things, such as a bunch of bananas or the teeth in the mouth; *pi7 is used for long things, such as a tail or a humming bird (because it has a long tail); and *7a7 is used for anything round, such as the moon, an eye, a turtle or the liver. In Káro, classifiers are used in possessive constructions, and following both a noun and its modifier in an NP. In (10) the classifier *pap 'long, big and round' is used with the noun 'alligator' and also following 'big':

(10) wayo nãk pap tãp pa6 alligator mouth CL big CL
'Alligator's big mouth'

5 PRONOUNS

All Tupí languages have pronominal prefixes to verbs. In most languages there can be just one pronominal prefix per verb (Mundurukú is an exception, allowing two). In the Tupí-Guaraní branch there are several prefix sets (one for A or S, one for O or S, etc.) – see §6 of chapter 5. It is probably the case that in all Tupí languages prefixes to nouns marking alienable and inalienable possessors are similar in form to the verbal prefix sets.

Crofts (1973: 83–94) lists four pronominal prefix paradigms for Mundurukú. Although the details are not totally clear it appears that one prefix set is used for the S argument of an intransitive verb of motion; a second set for the S argument of an active intransitive verb and the A of a transitive; a third set for the S argument of an adjectival-like stative verb (this is similar to the inalienable possessive prefixes on nouns) and the final set for subject of a limited set of verbs including 'know' (this is similar to the alienable possessive prefixes on nouns).

Languages from the Mondé branch (and at least some from the Tupari branch) include two verbal elements in each clause – a lexical verb, and an auxiliary element that marks tense, aspect and mood. The basic structure of a transitive clause is:

(a) either an NP or a pronominal prefix (to the following auxiliary)
realizing the A argument
(b) an auxiliary
(c) either an NP or a pronominal prefix (to the following verb)
realizing the O argument
(d) a verb.

In an intransitive clause, (a) marks the S argument and a pronominal prefix at (c) repeats this information (there are minor differences between the forms of prefixes in slots (a) and (c)).

We can illustrate this with three example sentences from Gavião (Moore 1984: 74, 80, 90). In (11) the A argument is realized as a pronoun and the O argument as an NP.

(11) maãa dza-bpiõO pogõ-å 1sgA + aux house-wall cover-BOUNDARY.MARKER
'I covered the walls.'
In (12) the S argument is realized as a pronoun in slot (a) and repeated in slot (c):

(12)  
\[
\text{dyáa} \quad \text{paa-gá-á} \\
\text{1pl.inclS + aux} \quad \text{1pl.inclS-go-boundary.marker} \\
\text{‘Let’s go!’}
\]

And in (13) the S argument is realized as an NP in slot (a) and as a pronoun in slot (c):

(13)  
\[
\text{[a-tsap kónj] dzáño mágá aa-kaá} \\
\text{3sgcoref-house to 1sg + brother 3sgS + aux 3sgS-go} \\
\text{‘My brother goes to his own house.’}
\]

(Note that in (13) the prefix a- to tsap ‘house’ cross-references dzáño ‘my brother’.)

Thus, in Gavião (and also in Suruí and Mekens) the pre-auxiliary element is A or S and the pre-verbal one is O or S (a mixture of accusative and ergative characteristics).

Table 4.5 shows the forms of pronominal prefixes in a selection of Tupí languages. It will be seen that 2sg is or begins with e- in each language except Karitiana where (as described in §2.1) e has shifted to a. 1sg exhibits more variety, being a- or a- or we- or wi- or a- (the in Karitiana is a development from a*). An original form we- could naturally have developed into u- or a- or we- (but not so plausibly to wi- or a-). 2pl begins with e- in most languages, suggesting that it was originally based on 2sg; the increments to e- vary considerably. 1pl.incl begins with o or a in five rows, followed by i or r or z or y, followed by o or a or e; it may be possible to relate these, once systematic phonological correspondences between the languages are worked out (paying attention to all parts of the grammar and lexicon). 1pl.excl shows the most variation, suggesting that the inclusive/exclusive contrast is a recent innovation, and has developed separately in each language (Xipáya is reported not to have this distinction). There is often a single 3rd person prefix, irrespective of number; the forms vary widely between languages.

6 VERBS

The languages vary a good deal in how they mark categories of aspect, tense and mood. As already mentioned, in the Mondé and Tupari branches these are suffixed to the auxiliary. The limited data available on Karitiana suggests that there are tense suffixes to the verb in a positive but not in a negative clause (Landin 1984). In Munduruku aspect and tense particles can either precede or follow the nucleus of the clause (predicate plus its core arguments), with aspect being nearest to the nucleus and tense further out (with a further peripheral constituent, marking location, etc., being further out than tense) (Crofts 1973: 54). Munduruku also uses reduplication to indicate present continuous, e.g. a’l’dzók’dzó ‘to be taking a bath’ (Crofts 1985: 68). Many languages appear to use the bare verb stem in imperatives.

Evidentiality is a pervasive feature of parts of Amazonia. On the information available, it is not a major characteristic of the Tupí family, being found in only a few Tupí languages. Suruí (W. Bontkes & Dooley 1985: 167-9) has a ‘hearsay’ particle, e.g.:

(14)  
\[\text{é Méresór iyá aka be kánê} \]
\[\text{DEM Name HEARSAY kill COMPLEMENTIZER want} \]
\[\text{‘He wanted to kill Meresor, I heard.’} \]

Suriá also has particles iná which ‘indicates factuality for a present event or state, or for a past event or state with present relevance’; and éniá, which ‘indicates factuality for a simple past event (whose present relevance is not being alluded to)’. (Evidentiality is also reported for Karitiana, Gavião and Káro.)

In Gavião there is marking on the verb for the number (singular or plural) of the S or O argument. A few verbs have suppletive forms, e.g. kaá ‘go (singular S)’ and máhá ‘go (plural S)’; other verbs use a suffix to mark plural S or O (Moore 1984: 158-9).

It appears that in Tupí languages there is a categorization of verbs as either strictly transitive or strictly intransitive. And there are, as would be expected, derivational processes for changing transitivity. Reflexive and reciprocal involve a verbal prefix (following the pronominal prefix) that derives an intransitive stem. Mawé has separate prefixes: -rëwe- for reflexive and -tòto- for reciprocal, e.g.:

(15)  
\[\text{[a-tsap kótj] dzáño mágá aa-kaá} \]
\[\text{3sgcoref-house to 1sg + brother 3sgS + aux 3sgS-go} \]
\[\text{‘My brother goes to his own house.’}
\]

\(\text{Table 4.5 Pronominal prefixes to the verb in some Tupí languages}\)

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2sg</th>
<th>1pl.incl</th>
<th>1pl.excl</th>
<th>2pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>i-</td>
<td>a-</td>
<td>iy-</td>
<td>ita-</td>
<td>ay-</td>
</tr>
<tr>
<td>II</td>
<td>a(y)-, (i()-</td>
<td>e(y)-</td>
<td>kás-</td>
<td>ti-</td>
<td>ozo-</td>
</tr>
<tr>
<td>III</td>
<td>u-</td>
<td>e-</td>
<td>se-</td>
<td>ule-</td>
<td>e-</td>
</tr>
<tr>
<td>IV</td>
<td>u-</td>
<td>e-</td>
<td>(w)a-</td>
<td>uru-</td>
<td>eIí-</td>
</tr>
<tr>
<td>V</td>
<td>Suruí</td>
<td>o-</td>
<td>e-</td>
<td>pa-</td>
<td>tóy-</td>
</tr>
<tr>
<td>VI</td>
<td>Munduruku</td>
<td>o'-, we'-</td>
<td>e'-</td>
<td>wiI(e')/a'-</td>
<td>o'we'-</td>
</tr>
<tr>
<td>VII</td>
<td>Káro</td>
<td>o-</td>
<td>e-</td>
<td>r'-</td>
<td>tI'-</td>
</tr>
<tr>
<td>VIII</td>
<td>Proto-Tupí-Guarani</td>
<td>a-, wi-</td>
<td>tfé-</td>
<td>né-</td>
<td>ya-(né/fe)-</td>
</tr>
</tbody>
</table>

(see chapter 5)
Aryon D. Rodrigues

(16)  
\[ \text{e-rew-kua} \]
\[ 2\text{sgS-REFL-know} \]
'You know yourself.'

In other languages a single prefix covers both reflexive and reciprocal, e.g. -dʒɛlwe' in Mundurukú (Crofts 1985: 187). A number of languages have a general intransitiveizing prefix, which covers reflexive, reciprocal and passive, e.g. -we- or -we- in Suruí (Van der Meer 1985: 228 and p.c.), -pe- in Káro and -be in Gavião (Moore 1994: 159).

In most (perhaps all) Tupí languages an intransitive verb can be made transitive by the addition of a causative prefix. This has the form -m- in Karitiána, -ma- in Gavião and Káro, -mo- in Aweti, -mu- in Tupí-Guaraní, -ma-, -mo- or -mu- in Mawé, and -mr- in Mundurukú. It can be exemplified for Káro:

(17)  
\[ \text{iyi! w-e-t a-ma-wiy-a} \]
\[ \text{squeeze Isg-Aux-PAST(?) 3sg-CAUS-go.out-PARTICIPLE} \]
'I squeezed it out (referring to a foot worm).'

8 DEMONSTRATIVES

Although there is only information on a few languages, it seems that the Tupí family is characterized by rich systems of demonstratives. For Suruí, Van der Meer (1985: 225) gives:

\[ \text{aáj 'this' (within speaker's reach, or in their hand)} \]
\[ \text{yá 'that' (close to hearer, or just out of easy reach of both speaker and hearer)} \]
\[ \text{amb 'away from both speaker and hearer)} \]

There are a number of other demonstratives in Suruí, including two that are size-related: aájín 'this small' and ayúnin 'that big'. (Moore 1984: 142 gives a similar system for Gavião.)

Suzuki (1997) devotes a whole MA thesis to discussing the set of fifteen or so deictics in Mawé, involving such parameters as 'visibility', 'gestural', 'distance', 'in direction of/ in trajectory of' and 'sitting/ standing/ scattered about'.

8 QUESTIONS

Polar questions are generally marked by a special particle. In Gavião the question particle te comes at the beginning of a clause (Moore 1984: 105) but in Mawé the particle apo comes at the end of the clause. Thus:

(18)  
\[ \text{ete [awiri tfjáde a-peyare ewe ikin o-or]} \]
\[ \text{THEN dog IMPERFV 3sg-bark COMPLEMENTIZER see 1sg-come é} \]
\[ \text{BOUNDARY.MARKER} \]
'Then I came and saw the dog barking.'

The sentence is, literally, 'Then the dog was barking, I came [and] saw [that]'.

9 SUBORDINATE CLAUSES

As in many Amazonian languages, subordinate clauses are often achieved through nominalization — see, for example, Moore (1989) on Gavião. Adverbial and complement clauses may be marked by special particles, such as be in (14) from Suruí.

Brandon and Seki (1981) provide a useful discussion of complementizers, including the following example from Mawé:

(19)  
\[ \text{kat pote it ere-to i lu-wiwo hap} \]
\[ \text{WHAT FOR NEG 2sg-go NEG 1sg-WITH NMLZR 1sg-know want} \]
'I want to know why you won't go with me.'

C. Bontkes (1985) describes the techniques for forming subordinate clauses in Suruí. The complementizer ewe, for instance, typically marks a complement clause in O function to 'see' or 'hear', as in:

(20)  
\[ \text{été [awiri tfjáde a-pekare ewe ikin o-or]} \]
\[ \text{THEN dog IMPERFV 3sg-bark COMPLEMENTIZER see 1sg-come é} \]
\[ \text{BOUNDARY.MARKER} \]
'Then I came and saw the dog barking.'

The sentence is, literally, 'Then the dog was barking, I came [and] saw [that]'.

10 PIVOTS

As we have seen, Tupí languages have some ergative characteristics — for instance, in some languages bound pronouns cover just S and O functions. There is a single mention in the literature of the syntactic orientation of a Tupí language. Van der Meer (1985: 210) suggests that Suruí has an S/A pivot (that is, it is accusative at the interclausal syntactic level). Compare:
(21) \(ók_{₁} \text{ déékaa, } \text{mekó}_₂ \text{ aka é man PERFV+go jaguar kill DECL.} \)

'The man went and the man killed the jaguar.'

(22) \(ók_{₁} \text{ déékaa, } \text{mekó}_₂ \text{ aka é man PERFV+go jaguar kill man kill DECL.} \)

'The man went and the jaguar killed the man.'

In (21) 'the man' is in pivot function in each clause (S in the first and A in the second) and its second occurrence is omitted. In (22) 'the man' is in a non-pivot function, O, in the second clause and thus cannot be omitted.

BIBLIOGRAPHY


**INTRODUCTION**

The Tupí-Guaraní branch (or subgroup) of the Tupí family is perhaps the best-known genetic grouping in Amazonia. It is named for the language groups which were most prominent during the period of the colonization of eastern South America: Tupinambá and Guaraní. The Tupinambá were coastal Indians, living mainly around the area where Rio de Janeiro is now located and northward, and were the main group with whom the Portuguese colonists had contact. They covered such an extensive area that their language was referred to in the seventeenth and eighteenth centuries as the ‘Brasilica’ language and as ‘Braziliano’, with the name Tupinambá appearing in the eighteenth century (Rodrigues 1986a). It was extensively documented by Padre Joseph de Anchieta (1595), who referred to it as ‘the most-used language on the Brazilian Coast’. This language, now extinct, has had a major influence on Brazilian Portuguese. According to Rodrigues, out of a list of 550 fish native to Brazil, 46 per cent have common names borrowed from Tupinambá. Out of 1,000 bird names, around 350 have Tupinambá origin. Many place names in Brazil also have Tupinambá origin, such as *lpiranga* ‘red water’ (from *y-pirag-a* ‘water-red-nominal suffix’). A few Tupinambá words have even entered Portuguese, such as *lpirangada* ‘red water’.

---

1 Extensive linguistic documentation has been done on various languages of the Tupí-Guaraní subgroup, the result of years of work by many linguists. The author thanks all of the linguists whose material is cited in this chapter for their contributions which have made so much reconstruction possible. Special thanks are given to those who contributed specific details to this chapter through personal communication: José Barrientos, Ana Suelly Alves, Wolf Dietrich, Bob Dooley, Carl Harrison, Perry Priest, Lucy Seki and especially Aryon Rodrigues. Rodrigues has spent over forty years studying the Tupí-Guaraní family, revising his phonological reconstructions, refining the genetic classification, and inspiring and orienting the present generation of Tupí-Guaraní comparativists, of whom I am one. He has also made accessible a wealth of historical information. Thanks also go out to Ruoleta Juliao, Kelly Smith, Helen Pease and La Vera Betts for supplying me with the specific locations of many of the Tupí-Guaraní languages.

2 The Tupinambá communities in Pará disappeared in the early eighteenth century. The Tupinikín (of Espirito Santo) and Potiguará (of Paraíba), the only remaining indigenous communities descended from the Tupinambá, speak only Portuguese (Rodrigues p.c.).
There are actually several languages which belong to the Guarani subgroup (i.e. the languages of subgroup 1, table 5.1) that are still spoken today by groups in southern Brazil, Bolivia, northern Argentina, and Paraguay. Some of these are referred to in their respective localities as Guarani (for example, Mbyá in Brazil and Chiriguano in Bolivia). It cannot be said with certainty that any of the varieties of Modern Guarani are a direct continuation of Old Guarani. However, the most likely candidate, other than Paraguayan Guarani, spoken extensively by a non-indigenous population, is Nhandéva.

Another group, the Tupi, whose language is closely related to Tupinambá yet lacks final consonants like Guarani, lived in what is now São Paulo, in the area of São Vincente and the upper Tieté River.4

All of these languages came to be used extensively among colonists in their respective regions. At the time of the colonization, the colonists were far outnumbered by the indigenous population. Since many of the colonists came to Brazil without wives, they ended up taking indigenous women as wives, with the result that the indigenous language became the mother tongue of their mestizo children.5 In this way the indigenous language began to be spoken in a completely new social context, becoming the predominant means of communication among the colonists and between them and indigenous groups. In the process of its expansion among the colonial population, the language underwent creolization, involving the progressive simplification of grammatical forms. In this way Tupi Austral (also called Língua Geral of São Paulo) grew out of Tupi in the São Paulo region. It was the predominant language of that area in the seventeenth century, but was replaced by Portuguese in the eighteenth. Xheengatú (Língua Geral of Amazônia), which means 'good talk', is still spoken today as the first language of a non-indigenous population in the upper Amazon region. It grew out of the Tupinambá language in the area of Maranhão and Pará and was the predominant language of communication by the Portuguese occupying the Amazon region during the seventeenth and eighteenth centuries.6 Paraguayan Guarani developed from Old Guarani - in a way that was heavily affected by Spanish - and is today one of the two national languages of Paraguay (Spanish and Paraguayan Guarani exert a considerable mutual influence on each other - see Dietrich 1993, 1995a).

4 In older writings Tupinambá has also been referred to as Tupi.
5 Padre Antônio Vieira wrote in 1694: 'It is true that today the families of the Portuguese and the Indians in São Paulo are so interconnected that ... the language spoken in these families is that of the Indians; and Portuguese is learned by the children at school' (Rodrigues 1986a).
6 In the eighteenth century the Portuguese government sent a royal decree prohibiting the use of the indigenous language by the Luso-Brazilian population (Rodrigues 1986a).
The use of the term Tupí-Guaraní preceded any attempts at linguistic classification. The term was used, for example, by Alfred Métraux, the first anthropologist to systematically study and make use of the data from chronicles of the 1500s and 1600s about Tupinambá and Guaraní. Although there were attempts at classification of the Tupi-Guarani languages in the 1940s and 1950s, the classification which Rodrigues (1958) presented to the 32nd International Congress of Americanists in Copenhagen in 1956 is the first to distinguish actual languages by the detailed use of linguistic data rather than to attempt classification based on a list of names and geographical factors. This classification also distinguished Tupi-Guarani languages from other Tupi languages with a high degree of accuracy, although data of many languages was insufficient at that time to permit much internal classification.

2 IDENTIFICATION OF TUPÍ-GUARANÍ LANGUAGES

Tupí-Guaraní languages are found throughout Brazil and beyond its borders, to northern Argentina (subgroup 1, see table 5.1), Paraguay (subgroup 1), Bolivia (subgroups 1 and 2), and French Guiana (subgroup 8). A few Tupi-Guaraní languages are now extinct, while others have been recently discovered. Initial contacts were made in the Brazilian state of Pará with speakers of two previously unknown language groups as recently as the late 1980s: Zoé (originally called Poturu) and another language simply referred to as 'the language of Aurá e Aufré', the two known speakers of that language.

The question of what is a language in the linguistic sense (as opposed to the political sense – see 'Conventions followed' above) is a vexed one for Tupi-Guarani. In subgroup 1, Avá and Izoceno are mutually intelligible and could be regarded as dialects of one language. The same applies to Kaiwá, Mbyá, Nhandeva and Paraguayan Guaraní. Guayakí and Xetá appear to be sufficiently different to be considered distinct languages. In subgroup 2, Sirionó/Yuqui and Jorá may comprise one language.

2.1 Distinguishing characteristics of Tupí-Guaraní languages

As one of the seven branches of the Tupi family, Tupí-Guaraní is noted for a high degree of lexical and morphological similarity among its member languages in spite of their extensive geographical separation.

Over forty languages or dialects have been identified as members of the Tupi-Guarani subgroup, in terms of the structural characteristics outlined in this chapter, and through having recurring lexemes such as forms phonologically derivable from:
### Table 5.1 Subgroups of Tupí-Guarani

<table>
<thead>
<tr>
<th>Subgroup 1 – Guarani subgroup (Argentina, Bolivia, Brazil, Paraguay)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chiriguano</strong> cluster</td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>Ává(^8)</td>
</tr>
<tr>
<td>1b</td>
<td>Izoteño</td>
</tr>
<tr>
<td>1c</td>
<td>Guayaki (Aché)</td>
</tr>
<tr>
<td>1d</td>
<td>Kawai (Pai-Tavyterê)</td>
</tr>
<tr>
<td>1e</td>
<td>Mbyá</td>
</tr>
<tr>
<td>1f</td>
<td>Nhändéva (also called Chiripá)</td>
</tr>
<tr>
<td>1g</td>
<td>Xeté</td>
</tr>
<tr>
<td>1h</td>
<td>Old Guarani</td>
</tr>
<tr>
<td>1i</td>
<td>Paraguayan Guarani</td>
</tr>
<tr>
<td><strong>Subgroup 2 (Bolivia)</strong></td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>Guarayu</td>
</tr>
<tr>
<td>2b</td>
<td>Sirionó</td>
</tr>
<tr>
<td>2c</td>
<td>Jorá (Hora)</td>
</tr>
<tr>
<td><strong>Subgroup 3 (Brazil)</strong></td>
<td></td>
</tr>
<tr>
<td>3a</td>
<td>Tupí</td>
</tr>
<tr>
<td>3b</td>
<td>Tupí Austral (Língua Geral of São Paulo)</td>
</tr>
<tr>
<td>3c</td>
<td>Tupinambá</td>
</tr>
<tr>
<td>3d</td>
<td>Nheengatá (Língua Geral of Amazônia)</td>
</tr>
<tr>
<td>3e</td>
<td>Kokáma/Ormáwa</td>
</tr>
<tr>
<td><strong>Subgroup 4 (Brazil)</strong></td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>Ává (Canoeiro)</td>
</tr>
<tr>
<td>4b</td>
<td>Asurini of Tocantins (or of Trucará)</td>
</tr>
<tr>
<td>4c</td>
<td>Suruí of Tocantins (or of Pará)</td>
</tr>
<tr>
<td>4d</td>
<td>Parakanã</td>
</tr>
<tr>
<td>4e</td>
<td>Tapirapé</td>
</tr>
<tr>
<td>4f</td>
<td>Guajajara</td>
</tr>
<tr>
<td>4g</td>
<td>Tembé</td>
</tr>
<tr>
<td><strong>Subgroup 5 (Brazil)</strong></td>
<td></td>
</tr>
<tr>
<td>5a</td>
<td>Araweté (?)(^9)</td>
</tr>
<tr>
<td>5b</td>
<td>Asurini of Xingu</td>
</tr>
<tr>
<td>5c</td>
<td>Kayabi</td>
</tr>
</tbody>
</table>

---

8 According to José Barrientos (p.c.) the speakers of this cluster in Bolivia (who number about 50,000) prefer to be called Guarani rather than Chiriguano. (There are indigenous groups in Brazil, as well, who are referred to by a name that is not their autodenomination.) However, since the name Guarani can be used to refer to any name in the subgroup, since ‘Bolivian Guarani’ is not spoken exclusively in Bolivia, and since the name Chiriguano has been used extensively in published sources, including by Dietrich (1986), it is used for the sake of clarity in this paper. The name Guarayu is used in Paraguay to refer to the Ává variety of Chiriguano, which is quite different from the Guarayu of Bolivia, placed in subgroup 2 here.


10 According to Cabral (1995) 98 per cent of the 19,000 members of the ethnic Kokáma group in Peru are bilingual, with no specific information available as to their level of fluency in either Kokáma or Spanish.

11 It is not yet possible to make a definitive conclusion as to the status of Araweté’s subclassification.
Table 5.2  **Proto-TG vowels**

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i</td>
<td>y (i)</td>
<td>u</td>
</tr>
<tr>
<td>low</td>
<td>e</td>
<td>a</td>
<td>o (i)</td>
</tr>
</tbody>
</table>

Details of the locations where the language is spoken and the approximate number of speakers\(^3\) are provided in the last column of the table.

Rodrigues based his subclassification primarily on phonological data, but recognized the need for the development of a more complete set of criteria, including grammatical and lexical features. The obvious difficulty is in acquiring sufficient complete information for each language to employ such criteria systematically. Dietrich (1990b) has worked on developing a set of morphological features in addition to a set of phonological features for establishing the linguistic proximity of languages. Mello (1992) is currently developing a lexical base for comparison.

### 3  **PROTO-TUPI-GUARANI PHONOLOGY**

#### 3.1  **Proto-Tupí-Guarani Phonemes**

Proto-Tupí-Guarani (proto-TG) had a predominantly CV syllable pattern in non-final syllables. The final syllable could be CV or CVC. Stress in the protolanguage, as in most of the descendent languages, occurred on the final syllable of the stem. Stress occurs on the penultimate syllable of stems in Chiriguano and the languages of subgroup 2 and of unsuffixed or uncliticized stems in Wayampi. In Asurini of Tocantins the stress is also on the penultimate syllable, except for nouns, which have stress on the third-to-last syllable, due to the permanent attachment of the (former) suffix -a (see §7.1) to nominal stems.

The proto-TG vocalic system had six vowels as shown in table 5.2. For two vowels a convenient orthographic symbol is used, with the phonetic value indicated in square brackets.

\(^3\) The principal sources of information used in the preparation of the population statistics in table 5.1 include Rodrigues (1984/5, 1986a) and Grimes (1996). Whenever possible, an attempt was made to use the most recent information, including personal communication and, for those Indians living in Pará, recent statistics in the *Liberal* newspaper. Where there is extensive use of the national language among an indigenous group, it is sometimes difficult to obtain accurate data regarding the number of speakers of the indigenous language.
Table 5.3  Proto-TG consonants

<table>
<thead>
<tr>
<th>VOICELESS STOP</th>
<th>LABIALIZED STOP</th>
<th>PALATALIZED STOP</th>
<th>VOICELESS AFFRICATE</th>
<th>FRICATIVE</th>
<th>NASAL</th>
<th>LABIALIZED NASAL</th>
<th>PALATALIZED NASAL</th>
<th>LIQUID</th>
<th>SEMIVOWEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p'</td>
<td>p*</td>
<td>p</td>
<td>3</td>
<td>m</td>
<td>m*</td>
<td>m(?)</td>
<td>r</td>
<td>w</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>3</td>
<td>n</td>
<td>g</td>
<td>g*</td>
<td>j</td>
<td>j</td>
</tr>
<tr>
<td>k</td>
<td>k*</td>
<td>k</td>
<td>k</td>
<td>z</td>
<td>g</td>
<td>g</td>
<td>g*</td>
<td>3</td>
<td>j</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3*</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Since only two vowel heights occur, the precise degree of height varies from language to language. Similarly, the high central vowel is more fronted in some languages than in others.

Stressed vowels (stem-final) may be oral or nasal, but nasalization is a property more of the morpheme than of the vowel and acts regressively.

The proto-TG system of consonants (Rodrigues and Dietrich 1997) is given in table 5.3.

The following consonants could occur word-finally: */3, *r, *k, *m, *n and *I.

The semivowel *j and possibly *w also occurred in this position. Medially only *j and perhaps *? occurred at the end of a syllable.

Most languages do not permit *j word-initially, and this was a possible restriction in proto-TG as well.

The language which shows the least change from proto-TG, both phonologically and morphologically, is Tupinambá.

3.2  Proto-Tupí-Guaraní allomorphs

The most salient cases of allophonic variation were those of the nasal consonants, which were pure nasals ([m], [n], [g]) in a nasal environment (preceding a nasal consonant or a nasalized vowel) and post-occluded nasals ([mb], [nd], [g]) in an oral environment.

3.3  Possible Tupí-Guaraní morphophonemic phenomena

Several likely morphophonemic rules for Tupí-Guaraní languages are discussed in detail in Jensen (1989). In this chapter I will only mention a couple of types.

3.3.1  Replacement by nasal consonants

(i) Progressive application

A nasalized vowel (1) or a nasal (2) causes the initial voiceless consonant of the following morpheme in the same phonological word to be replaced by a corresponding nasal. For example, from Tupinambá:

(1)  nupá 'beat' = katá 'good' → nupáŋatú 'beat hard'
(2)   moh 'CAUS' + só 'go' → monó 'send (cause to go)'

(ii) Regressive application

A nasal consonant has the effect of optionally replacing by a nasal, a voiced consonant that precedes it in the same word, as in the following example from Tupinambá:

(3)  ero 'COMIT.CAUS' + sém 'go out (exit)' → enosém 'to go out, causing O to go out as well'

Evidence in several languages supports the existence of both of these rules in the protolanguage. Even in languages where they are no longer productive, numerous fixed forms still occur, for example, mono and ono in Wayampi, even though the only productive nasalization is the type in (1).

3.3.2  Evidence from Tupí-Guaraní languages

Some morphophonemic rules depend on the type of morpheme juncture that occurs. A juncture of two stems is indicated by =. A juncture of stem and affix or of two affixes is indicated by +.
3.3.2 Strategies for the loss of consonant clusters formed at morpheme juncture

(i) Epenthesis

When a consonant-final morpheme combines with another morpheme, various morphophonemic strategies may apply, depending on the nature of the second morpheme. If the second morpheme is a consonant-initial suffix, an epenthetic vowel, \( y \) or \( i \), is inserted. For example, in the Amapari dialect of Wayampi: *wyr 'underside' + pe 'PUNCT.LOC' > wyrype 'under'. In Kayabí: *mytér 'middle' + pe > myteripe 'in the middle'. The environment in which epenthesis takes place is limited since several suffixes have vowel-initial allomorphs for consonant-final stems, alternating with consonant-initial allomorphs for vowel-final stems. For example: *-watju ~ *-utju 'large'.

(ii) Loss

When both morphemes are stems and the second morpheme is consonant-initial, the final consonant of the first morpheme is lost: *tarey 'traíra fish' = *tío 'white' > *tareytió 'white traíra (species)'. If the final consonant of the first stem is a nasal, the feature of nasalization remains on the preceding vowel, as in Tupinambá:

(4) akáj 'head' = póy 'flat' → akáj = méy (through rule in §3.3.1) → akäméy 'flat head'

When the initial consonant of the second morpheme, whether stem or suffix, begins with a glottal stop, in most languages the glottal stop is lost and the final consonant of the preceding stem is retained, as in Wayampi:

(5) *akáj 'head' = *ʔók 'remove' → akáj 'decapitate'
*taʔár 'bamboo species' + *ʔi 'DIM' → takári 'bamboo species'

This inversion of loss strategy can be explained by a rule of metathesis prior to loss. In Parintintín and Kayabí this metathesis occurs without loss of the other consonant.

(6) (P) tiý 'white' + ʔi 'DIM' → tiʔií 'very white'

(7) (Kb) ipit 'skin' = ʔok 'remove' → ipiʔok 'to skin'\(^{16}\)

The occurrence of syllable-final glottal stop as the result of morpheme combination in these two languages suggests that the same may have occurred in the proto-language.

\(^{16}\) See §4.5.1 (on devoicing) regarding alternation of \( t \) and \( r \).

## 4. Phonological Changes within Tupí-Guaraní

The phonological diversification of Tupí-Guaraní languages took place primarily through weakening of *\( p \), *\( t \) and *\( ts \); through various mergers, complete or partial; through other changes in manner and point of articulation which did not result in merger; and through the partial or complete loss of final consonants. Vocalic shift, loss of nasalization, and stress change also contributed to this diversification. The consequences of these changes can be lexical, as in the creation of homonyms, or morphological. Some of these changes can be seen in a general way in table 5.4.\(^{17}\)

### 4.1 Weakening of *\( t \) and *\( ts \)

The weakening of *\( t \) and of *\( ts \) follows a path commonly attested in languages: *\( t \sim t > ts \) or *\( f \sim f > h > a \). The only subgroup in which the distinction between these

---

\(^{17}\) For a more detailed account of changes, language by language, see Dietrich (1990b).
two phonemes is systematically retained in subgroup 1, as in the Mbyá reflexes of *tsó ‘mother’ and *tsó ‘(to) go’: tʃá and ʃ. In other subgroups the two phonemes have merged. For example, the Tupinambá forms for these two morphemes are sʃ and sð, in Guajajara they are hʃ and hð, and in Wayampi ʃ and ʃ. Nevertheless, even in languages where merger has taken place, there are indications that *ts in certain environments (or certain morphemes) was more resistant to weakening than was *ʃ. For example, in Wayampi, where normally *ʃ and *s became o, the reflex of *ʃuʔu ‘(to) bite, chew’ is sʔu, and in Asurini of Tocantins, where normally *ʃ and *s became o or a, the reflex of the first person singular pronoun *ʃó is sð.

Comparative data show that h is particularly resistant to loss in the final, (originally-) stressed syllable. For example, the h is lost in initial (prestressed) position (10) but retained in stressed position (11) both in Parintintí and in Asurini of Tocantins.

(10) *tsó ‘(to) go’> hó (P), há (A)
(11) *tsó ‘(to) go’> hó (P), há (A)

Medially in (originally-) prestressed syllables, the h is retained in Parintintí and lost in Asurini, as in example (12).18

(12) *tskyjé ‘fear’> k̪tskyjé > khyjí (P), kþýse (A)

The h is also resistant to loss in the monophonemic person marker for third person: *ts: 3 (allomorph for Class II stems). Presumably there is resistance to the formation of a o morpheme. Hence, h is retained in the words h-orywète (A) and h-irywété (Gj) ‘(to be) happy’ (<*orẙ̊̊s-été ‘happy-genuine’), even though the same phoneme was lost in Asurini in the word-initial position in the reflex of *sʃeši ‘worm’: éwóʔia. In Wayampi, the reflexes of both *ʃ and *s were lost, leaving a o morpheme as the reflex of *ts-, as in o-á ‘his/her eye’ (<*ts-éti), except for when it combines with a few monosyllabic stems in the Jari dialect, such as *ts-é ‘3-name’> h-ce.

4.2 Palatalization

Various phonological changes, either widespread or localized, in some way involved palatalization.19 In many languages the protophoneme *p, which occurs in the reconstruction *epík ‘(to) see’, underwent an unconditioned change of point of articulation to the alveo-palatal position, followed by weakening, as in example (13).

(13) *p > tʃ > f or ts > s20

In Tupinambá this consonant underwent a structural reanalysis as a consonant plus a semivowel, as in epík ‘(to) see’.

The phoneme *k occurs in the reconstructed morphemes *ik’é ‘(to enter)’, *k’é ‘(to sleep)’ and *k’é ‘here, near the speaker’, which were previously reconstructed as *iké, *kér and *ké, respectively. This change in analysis accounts for the palatalized form ké in Hoeller’s data (1932) of Guarayú.21 In Kayabi this phoneme resulted in the alveolar fricative s ‘set’(to) sleep’ and se ‘(to) enter’. In Guajajara it resulted in the alveo-palatal affricate in išé [iʃe] ‘(to) enter’; that this was caused by the palatalized consonant rather than the preceding high vowel can be seen in ikó (<*ikó ‘to be in motion’).22 In Parintintí *k provoked the raising of the subsequent vowel, then merged with *k, resulting in kir ‘sleep’ and kí ‘here’. In most other languages, generally *k merged with *k (and *g with *g).23

Other palatalized forms developed in various languages as allophonic variation. The palatalization of t before i, as in example (14), occurred in most languages, but not in Tupinambá or Parintintí.

(14) *t (preceding i) > tʃ > f or ts > s

The palatalization of k (including the originally labialized and palatalized phonemes) in Urubú-Kaapor contiguous to i resulted in i. In Tupinambá the fricative s (<*ts, *tʃ) was palatalized as f contiguous to i. In Guajajara the same sort of palatalization takes place, though the sources of s in this language are different from those in Tupinambá.

Some of these changes are limited to allophonic variation, such as the case of Tupinambá. However, what begins as allophonic variation has the possibility of leading to a phonemic split, as the palatalized allophone merges with another phoneme or with the palatalized allophone of another phoneme. Various combinations of mergers took place in the Tupí-Guarani subgroup as the reflexes of several palatalized consonants or of allophones of consonants converged in the alveolar and alveo-palatal positions.

In Guajajara the phoneme s, with allophone tʃ contiguous to i, has several sources: *p (> epík > ēsša), *t (> *ōatí > awasi [awatfj]) and *k (> *iké > išé).

In Urubú-Kaapor (Kakumasu 1986) there are cases of ʃ which are clearly

18 Metathesis occurred in this word in languages from subgroups I and 4-8.
19 I am using the term broadly to refer to changes which involved palatalization in some way, even if the final result was not a palatalized phoneme, such as s.
20 Anambé has h as the end result of this change, one step beyond s in the weakening process *epík ‘(to) see’> ahá (Juliano 1993).
22 The morpheme k’é ‘(to sleep)’ is kér in Guajajara, even though presumably at one time the reflex of *kéér would have been phonetically [tʃer].
23 In Kayabi the nasalized equivalent of s (<*k) is g: mọp ‘cause to enter’.
recognizable as having originated from *k because of an alternation in the initial consonant of the stem when it is preceded by the third person prefix i. For example: *kwd' < *k’rd' 'hole', i- *surd > its hole', kyr < *ky?l 'louse, lice', i-sy 'his louse, lice'. Such alternation is not observable medially, although the conditioning environment is still intact: i/kö 'to enter' (< *ik’ö), ijo 'continuative aspect, in motion' (< *i’ko).

Other sources of i come from the palatalization of t before *awa’i 'corn' (< *awa’i). At times the conditioning environment has been absorbed into the fricative, permitting alternations such as *fa?w and i *fa?w (< *i-káy’i) 'his bone'. Other morphemes have the appropriate conditioning environment but are clearly not of Tupi-Guaraní origin, such as pati (from Portuguese pásar 'to take a walk or a trip'). Thus what began in large part as palatalized allophones of two separate phonemes developed into a new phoneme in Urubú-Kaapor. Similarly, the reflexes of *j (to) kill' s6ka).24 To these were added the two forms of the first person singular pronoun (*juká > sil JJiman).

(kamísa 'cloth' < Portuguese camisa 'shirt'; semi ‘manioc grater' < Apalaí kie (see §3.3.1) because of an alternation in the initial con­

borders, so there is no reason to distinguish a set of labialized consonants from these sequences.

The distinction between the reflexes of *p and *k is retained in Kayabi and Kamayurá as well. However the reflexes themselves have changed to f [j] in Kayabi and hw in Kamayurá. A parallel change also took place in these two languages with the sequence *pu, which became ho in Kamayurá, as in *pu?á 'laugh' > huká, and fu in Kayabi, as in *pu?á 'laugh' > fuká. If *p were reinterpreted as a sequence in these languages, as in Tupinambá, a single rule could account for the change before w and u. This parallel change did not take place in the languages in which *p merged with *k.

4.4 Merger of *β with *w

In subgroups 4, 5, and 7, with the exception of the Amapari dialect of Wayampi (subgroup 8), there was a merger of *β with *w. This was a complete merger in the other languages of subgroup 8 (although no reflex of *β occurs word-finally) as well as in the Tenetehára cluster of subgroup 4. Examples (17) and (18) show this merger in syllable-initial and syllable-final positions, respectively.

(17) *yβytu 'wind' > ywytú (Gj, U), ywytu (WJ), but yβytu (WA)

24 The nasal equivalent of hw (see §3.3.1) is m as in *emi-paq 'person who receives orders' > eminaj, whereas *paj > haj (Seki p.c.). In Kayabi the nasal counterpart of f [j] is m and of k is n (Dobson 1988: 136).

(5) Tupi-Guaraní

proto-TG

Mbyá (subgrp 1)  ká k’erá k’á
Guaraní (2)  ká k’erá k’á
Guajará (4)  k’á k’eráw k’á
Parintintin (6)  k’á k’eráβ k’á
Wayampi (WJ) (8)  õk’a poéra k’a

In Tupinambá, the reflex of *p is unchanged except for a structural reanalysis as a sequence instead of a labialized consonant, as in example (16). In this language a whole series of sequences of consonant plus semivowel are created at morpheme boundaries, so there is no reason to distinguish a set of labialized consonants from

(16) proto-TG

Tupinambá (3) pwár pweráJ3 kwár
Kamayurá (7) hwát hweráp .kwat

For example:

as the sequence instead of a labialized consonant, as in example (15). In most of these languages the merger was complete, but in Wayampi it was limited to what were originally stressed syllables.

The data in Lemle (1971) does not reflect the merger of *j in Asurini of Tocantins with these other sources.

25 According to Seki (p.c.) this morpheme does not occur in isolation in Kamayurá, but occurs in the compound word ywy- kwá 'hole in the ground'.

26 The nasal equivalent of hw (see §3.3.1) is m as in *emi-paq 'person who receives orders' > eminaj, whereas *paj > haj (Seki p.c.). In Kayabi the nasal counterpart of f [j] is m and of k is n (Dobson 1988: 136).
In other languages the merger was incomplete, being limited to syllable-initial position, as in example (19). In syllable-final position (which in fact only occurs word-finally), the reflex of final */β/ is *w in Asurini of Tocantins (20). There was a primary split in languages of subgroups 5 and 7 (Kayabí and Kamayurá), with the reflex of final */β/ normally becoming reassociated with the *p phoneme.

(19) *yβytu 'wind' > ywytoa (A), ywyto (Km)
(20) *-paβ 'completely' > -pam (A), -pap (Kb)

4.5 Final-consonant phenomena

Various phonological phenomena deal specifically with word-final consonants.

4.5.1 Devoicing

For Tupinambá, Rodrigues describes an optional rule by which *β and *r in final position become voiceless unreleased stops, *p and *t respectively (Jensen 1989: 53–4). For example:

(21) sjé rúβ 'my father' > férúp
sjé ra?yr 'my son' > férúu t'j

This devoicing apparently became obligatory in the development of such languages as Kayabí, Kamayurá and Tapirapé. Whereas in Tupinambá the unreleased bilabial stop is easily recognized as an allophone of the bilabial fricative, such an analysis is more complicated in the other three languages due to the merger of *β with *w.

Linguists working in these languages have apparently opted for an analysis of a split, the [fβ] allophone merging with w and the [pτ] allophone with r, even though the alternation of p with w still shows up upon morpheme combination, as in the following example (22) of the morpheme */uβyβ 'arrow'/ from Kayabí, without and with the addition of the nominal suffix -a.

(22) uβyp 'arrow'
tapyyja ruβyw-a 'white man's arrow' (i.e. gun)

By analogy [t'] was treated as an allophone of t and not of r, even though the same type of alternation occurs with r on morpheme combination, as can be seen in the following example of the morpheme */aβyτ 'son, child'/.

(23) taβyt 'son'
ku?i raβyτ-a 'monkey child'

4.5.2 Nasalization

In Asurini of Tocantins all final consonants became corresponding nasals, as in the following pairs of unsuffixed verbs, which terminate in bilabials (24), alveolars (25) and velars (26).

(24) *pβ 'finish' > pam
tσem 'leave' > hem

(25) *kβr 'sleep' > ken
etún 'smell' > éton

(26) *?ök 'remove' > ?aq
pirág 'be red' > pirág

When the stems which terminate with formerly non-nasal voiced consonants (*β, *r) combine with another morpheme – for example, the oblique-topicalized suffix -i (see §9.2) – the non-nasal consonant is recuperated as w (due to the merger of *β with *w) or r. i-paw-i and i-kεr-i, respectively. The final *k, which was the only voiceless consonant which occurred in final position in proto-TG, is not recuperated in this environment: i-haw-i. Nouns originally terminating in *k also show a permanent change to *a: *mannotiök-a > manótau 'manioc'.

Nasalization at a word boundary is in fact not uncommon in the languages of lowland South America and can be accounted for phonetically by the fact that the normal resting (and breathing) position of the velic is lowered, as it is for nasals (Rodrigues 1986b). The nasalization of *k, i.e. the replacement by its correspondent nasal, occurred in a few words in the Amapari dialect of Wayampi: pάp ( <*pák) 'paca', yóq ( <*yfók) 'larva'. In Tapirapé there is optional nasalization of final *r.

4.5.3 Loss

In a number of languages, final consonants have been completely or partially lost, as in the following examples (27)-(28) from Mbyá, of subgroup I. When the final consonant was a nasal, the feature of nasalization is retained in the final vowel (28).

(27) *akóβ 'be hot' > akú
k*ár 'hole' > k*á
*pytsók 'to grasp' > pyý

There are other languages in which devoicing does not occur at all, such as Guajajara. In this language there was a complete merger of *β with *w, and the final consonants which come from *β and *r are phonemically and phonetically w and r, respectively, as in the agentive and circumstantial nominalizers, -har (< *-tsdr) and -awa (< *-tsdB), respectively.
(28) *kám 'breast' > kā
*pytún 'night' > pytū
*kān 'bone' > kā

Loss of final consonants is also a characteristic of subgroup 2, as in the following example (29) from Guarayu.

(29) *eymá(! 'domestic animal' > eyma
*pysyk '(to) grasp' > pysy
*tfám 'strillg, cord' > tsā
*kjér '(to) sleep' > ke

However, in this language the syllable-final r is retained in nouns (30). According to Newton (p.c.) the nominal form of 'sleep' retains the final r, which the verbal form does not.

(30) *kwar 'hole' > kwar

In subgroup 8 there is some variation as to the loss or retention of final consonants. In Urubú-Kaapor all final consonants are retained with the exception of *β. In Wayampi all final consonants have been lost in the Jari dialect. In the Amapari dialect a few nouns retain final n or g and several, but not all, retain r. Comparative data of Wayampi from a century ago by the French explorer Coudreau show the loss of final consonants to be a relatively recent phenomenon, since he registers the presence of all six final consonants (Jensen 1989).

The loss of final consonants has contributed towards the formation of homonyms, including the nominalizations of circumstance *-tsdr and of agent *-tsb: both have become -há in Kaiwá and -a in the Jari dialect of Wayampi.

4.6 Vowel shift

Languages of subgroup 4 are characterized by vowel shift in central and back vowels (in the general direction of u>a>y) and by loss of nasalization in some languages. Soares and Leite (1991) describe the two types of changes as interacting to produce the present vowel systems in these languages through a redistribution of allophones, which could have happened simultaneously for the various vowels. This is an alternative to the already-existing hypothesis of a push-chain of gradual changes.

Guajajára has a seven-vowel system and no nasalization. The nasal allophone of *a, being higher in its point of articulation than its oral counterpart, retained this position with the loss of nasalization and developed into a distinct phoneme á. In some environments (prestressed) the reflex of *o was raised to u.

Asuriní of Tocantins has a five-vowel system and no nasalization. The reflex of *o in the originally stressed (i.e. final) position changed to a in an (originally) oral environment. (Where *o occurred in both the final and the penultimate syllables, both occurrences underwent the same change.) The reflex of *u changed in all environments to o. The reflex of originally stressed *a in an (originally) nasalized environment usually became o, but occasionally became y.

Compare the data from these languages:

(31) *proto-TG Guajajára Asuriní

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>*kutúk</td>
<td>kutúk</td>
<td>kótōŋ</td>
</tr>
<tr>
<td>*monó</td>
<td>monó</td>
<td>máňa</td>
</tr>
<tr>
<td>*óré</td>
<td>uré</td>
<td>óře</td>
</tr>
<tr>
<td>*piráŋ</td>
<td>piráŋ</td>
<td>piroŋ</td>
</tr>
<tr>
<td>*kujá</td>
<td>kuzá</td>
<td>kóso</td>
</tr>
<tr>
<td>*akáŋ</td>
<td>okáŋ</td>
<td>ąkýo[-a]</td>
</tr>
<tr>
<td>*ti</td>
<td>si</td>
<td>si</td>
</tr>
</tbody>
</table>

In Tapirapé the changes *o>a and *u>o were both general. Nasalization was not lost. The nasalized phoneme *á became ą and the (originally) oral *a became nasalized á (phonetically a nasalized [a]).

Vowel shift is also described for Anambé of subgroup 8, although only for the final (stressed) syllable. Unlike the languages of subgroup 4, its vowel shift involves (formerly) low front and central vowels rather than (formerly) low central and back ones. The changes in central vowels are similar to Tapirapé, with the nasalized allophone of *a raising to ą and merging with *y, and the oral allophone becoming nasalized: *u>a. However, in this language the front vowel rather than the back one shifted to a. Compare cognates in Anambé with those of its relative Wayampi (32).

(32) *proto-TG Anambé Wayampi (WJ)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>*piráŋ</td>
<td>pirý</td>
<td>pirá</td>
</tr>
<tr>
<td>*pirá-čr</td>
<td>pirá</td>
<td>pire</td>
</tr>
<tr>
<td>*pirá</td>
<td>pirá</td>
<td>pira</td>
</tr>
<tr>
<td>*pi?tů</td>
<td>pi?ů</td>
<td>pi?ů</td>
</tr>
<tr>
<td>*nupá</td>
<td>nupý</td>
<td>númer</td>
</tr>
</tbody>
</table>
5 STEMS

5.1 Categories of stems

There are three basic categories of stems which can be inflected: nouns (§7), postpositions (§8) and verbs (§9). Verbs are subclassified as transitive, active intransitive and stative intransitive. Other categories of noninflected stems include quantifiers, demonstratives and time words. Adjectives do not exist as a separate category. Information which in many languages is communicated through adjectives is communicated in Tupí-Guaraní languages on a syntactic level through nouns in appositional phrases or through stative intransitive verbs. Such information is also communicated on the morphological level through suffixation or through composition (§7.2).

5.2 Stem classes

Inflected stems, regardless of their grammatical categorization, are divided into two arbitrary classes based on whether or not they combine with a linking or relational prefix *r-. This prefix, which occurs only with class II morphemes, indicates a tight grammatical relationship between the noun or person marker from set 2 which precedes it and the morpheme to which it is prefixed. The same relationship is unexpressed in class 1 morphemes. The following examples in (33) show the combination of class I without, and class II with, the r- morpheme in nouns, transitive verbs (O), stative intransitive verbs (S), and postpositions, respectively. In (33) the preceding morpheme is a person marker; in (34) it is a noun.

(33) Class I

*tfé akáŋ ‘my head’
*tfé potáŋ ‘like me’
*tfé katú ‘I am good’
*tfé koty ‘toward me’

(34) Class II

*pírâ akáŋ ‘fish’s head’
*pírâ r-etsá ‘fish’s eye’

6 PERSON MARKERS

Four sets of person markers have been reconstructed for proto-Tupí-Guaraní, as shown in table 5.5. Their function is discussed in more detail in §7–§9.

Set 1 markers are used with transitive and active intransitive verbs and refer to A

<table>
<thead>
<tr>
<th>Class</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
<th>Set 4</th>
<th>Free pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>a-</td>
<td>tfé</td>
<td>wi-</td>
<td>itfé</td>
<td></td>
</tr>
<tr>
<td>1excl</td>
<td>oro-</td>
<td>oré</td>
<td>oro-</td>
<td>oré</td>
<td></td>
</tr>
<tr>
<td>1incl</td>
<td>ja-27</td>
<td>jané</td>
<td>jere-</td>
<td>jané</td>
<td></td>
</tr>
<tr>
<td>2sg</td>
<td>ere-</td>
<td>né</td>
<td>e-</td>
<td>oro-</td>
<td>ené</td>
</tr>
<tr>
<td>2pl</td>
<td>pe-</td>
<td>pé</td>
<td>peje-</td>
<td>opo-</td>
<td>pe ... e</td>
</tr>
<tr>
<td>3</td>
<td>o-</td>
<td>-i-</td>
<td>-ts-</td>
<td>-t-</td>
<td>o-</td>
</tr>
</tbody>
</table>

and So, respectively. In most languages their use is restricted to independent clauses.

Set 2 markers are used to indicate the genitive in nouns and the object of postpositions, except when their referents are coreferential with the subject (A or S) of the clause. They also occur with verbs in the following constructions:

transitive verbs, referring to O, in subordinate, serial and oblique-topicalized constructions, and also in independent constructions when O is superior to A in a 1 > 2 > 3 hierarchy,

active intransitive verbs, referring to So, in subordinate and oblique-topicalized constructions;

stative intransitive verbs, referring to So, in all constructions.

As with nouns and postpositions, the verbs also carry coreferentiality restrictions. When first and second person markers occur with class II stems, they require the presence of the r- prefix on the stem. This prefix has an allomorph n- for second person plural. Note that the five forms for first and second person are the same as, or derived from, the free pronouns. The third person markers are prefixes: i- occurring with class I stems, ts- occurring with most class II stems, and t- occurring with a subset of class II stems.

Set 3 markers are coreferential markers, occurring with nouns, postpositions and verbs (particularly intransitive serial verbs), substituting the markers in set 2, when

Some languages have a prefix based on *ti- which occurs with transitive verbs. See Jensen (1987, 1998b) for a proposal of its derivation.

Rodrigues and Dietrich (1997) gives a different reconstructed form *jare-.

In some languages their use has been extended to other (non-independent) clause types. See Jensen (1990).

Details of this split-ergative marking system are discussed more fully in §9. This system has been replaced in Urubú-Kaapor by strictly nominative marking.

In several languages, including Asurini of Tocantins, Guajajara and Wayampi, first and second person markers of set 2 are presently analysed as prefixes, though there are strong reasons to consider that they were originally independent words.

27 Some languages have a prefix based on *ti- which occurs with transitive verbs. See Jensen (1987, 1998b) for a proposal of its derivation.
28 Rodrigues and Dietrich (1997) gives a different reconstructed form *jare-.
29 In some languages their use has been extended to other (non-independent) clause types. See Jensen (1990).
30 Details of this split-ergative marking system are discussed more fully in §9. This system has been replaced in Urubú-Kaapor by strictly nominative marking.
31 In several languages, including Asurini of Tocantins, Guajajara and Wayampi, first and second person markers of set 2 are presently analysed as prefixes, though there are strong reasons to consider that they were originally independent words.

...
their referents are the same as the subject (A or S) of the independent verb (or, in the case of nouns and postpositions, the dominating verb of the clause they are in).32

Set 4 markers occur with transitive verbs and are portmanteau forms indicating that A is first person and O is second person. The two forms *oro- and *opo- distinguish the number of the second person O, singular and plural, respectively, but do not distinguish the number of the first person.

The basic divergencies from the protosystem which may be observed in the various languages are the following:

- an extension of the use of set 1 forms to subordinate clause types.
- a reduction of the number of coreferential forms to only third person, and their replacement by other forms (set 1 or set 2) in certain constructions.
- various types of changes in set 4 morphemes, including their loss, their merging to a single form, oro-, and the occurrence of a form for second-person plural not directly derivable (phonologically) from *opo-.

In Parintintí and Kayabi the use of the reflexes of *i- and *ts- for third person (set 2) has been restricted to non-human referents. For human referents these two languages have three forms, specific for number and gender (in the singular), which are not found in the other subgroups of Tupi-Guaraní:

(35) Parintintí Kayabi (men's speech)

<table>
<thead>
<tr>
<th>Case</th>
<th>Parintintí</th>
<th>Kayabi</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg.masc</td>
<td>ga</td>
<td>?ga</td>
</tr>
<tr>
<td>3sg.fem</td>
<td>hé</td>
<td>éé</td>
</tr>
<tr>
<td>3pl</td>
<td>ña</td>
<td>?ñá</td>
</tr>
</tbody>
</table>

7 NOUNS

7.1 Case marking

In proto-TG, as demonstrated in many languages of the branch, any noun terminating in a consonant received a nominal case suffix *-a, whenever it occurred in a subject, object or genitive function (36)–(38). Nouns terminating in a vowel did not combine with this suffix.

There are four oblique cases: attributive case *-amo (C__)~ramo (V__) and three locative cases: diffuse (spread out in a certain area) *-bó, punctual (at some specific point) *-pe, and partitive (at some specific part of the whole, such as a body part) *-i. These are illustrated in Tupinambá in examples (36)–(39), respectively.

(36) e-i-mome?ú Tupá-ó r-a?yr-amo né
imper.3-tell God-nom.case link-son-attrib 2sg
r-ekó-ó
link-be-nmlzr
"Tell about your being God's Son.'

(37) ka?á-bó jwár-a ... r-ekó-w
jungle-dif.loc jaguar-nom.case link-be-obl.top
"Jaguars live throughout the jungle.'

(38) jwár-a sjé kó-pe s-ekó-w
jaguar-nom.case 1sg garden-punct.loc 3-be-obl.top
"The jaguar is in my garden.'

(39) sjé júr-i a-rekó
1sg neck-part.loc 1sg-have
'I have it at my neck.'

In some present-day languages, such as Asurini of Tocantins, the use of the nominal suffix has been extended to all nouns, whereas in other languages, such as Guajajara, its occurrence has been completely lost. Some of these locative cases have combined with nouns indicating spatial relationships to create words similar to postpositions. For example, in the Amapari dialect of Wayampi:

(40) *wýr 'underside' + *-pe 'punct(loc)'>wýrype 'under'
*?ár 'top' + *-bó 'dif(loc)>?arýbó 'above'
*pýr 'part next to' + *-i 'part(loc)'>pýri 'next to'
*pýr 'part next to' + *-bó 'dif(loc)'>pýrybó 'near, along'

Kayabi distinguishes punctual and diffuse forms of 'under': ?wýrype and ?wýrimó.

Nouns can also function as the predicate without any derivational marking other than the absence of *-a. In this case they are inflected in the same way as stative intransitive verbs. For example:

(41) *i-memýr-a 'her child'
*i-memýr 'she gave birth'
7.2 Possession

Pronominal possession is normally expressed by the person markers of set 2. The forms for first and second person were derived from free forms and are reconstructed for proto-TG as free forms. However, in many languages they are now analysed as prefixes. These forms co-occur with the prefix r- when combining with a stem from class II (43). The third person marker, undisputably a prefix, occurs without the r- prefix.

(42)

<table>
<thead>
<tr>
<th></th>
<th>Tupinambá</th>
<th>Guajajára</th>
<th>Wayampi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>*tjé pó</td>
<td>sjé pó</td>
<td>he-pó</td>
</tr>
<tr>
<td>1excl</td>
<td>*oré pó</td>
<td>oré pó</td>
<td>ure-pó</td>
</tr>
<tr>
<td>1incl</td>
<td>*jané pó</td>
<td>jané pó</td>
<td>zane-pó</td>
</tr>
<tr>
<td>2sg</td>
<td>*né pó</td>
<td>né pó</td>
<td>ne-pó</td>
</tr>
<tr>
<td>2pl</td>
<td>*pê pó</td>
<td>pê pó</td>
<td>pe-pó</td>
</tr>
<tr>
<td>3</td>
<td>*t-pó</td>
<td>i-pó</td>
<td>l-po</td>
</tr>
</tbody>
</table>

When the possessor of the noun is identical with the subject, a coreferential prefix (set 3) is used, as can be seen by comparing Tapirapé (44), where the arguments are not coreferential, and (45), where the arguments are coreferential.

(44)

erc-ma-pén fe-pá
2sg-caus-break 1sg-hand
"You broke my hand."

(45)

ää-ma-pén we-pá
1sg-caus-break 1sg-hand
"I broke my (own) hand."

In some languages, like Asurini of Tocantins and Tapirapé, these prefixes occur for all persons. In others, like Guajajára, the Guarani languages and Wayampi, they occur with just third person, the only person for which there is any possible ambiguity. In Wayampi, there is only one form for 'my hand': é-po, whereas there are two forms, non-coreferential and coreferential for the third person reference 'his/her hand': k-po and é-po, respectively. Urubú-Kaapor has lost the coreferential prefixes altogether.

7.3 Noun composition

There are a number of Tupi-Guarani suffixes which occur with nouns to give modifying information which in other languages tends to be communicated through a separate class of adjectives. These suffixes indicate:

- size *-watfú (and its allomorph *-utfú) 'large', *-tí 'small'
- quality *-êté 'genuine', *-rán 'imitation'
- status *-pér (and its allomorphs *-wér and *-éär) 'former', *-rám (alternating with *-wám and *-ám) 'future'

For example, from Tupinambá: *j-wasú 'great water', pajé-rám-a 'future shaman'.

Nouns may also occur in composition with verbs, particularly stative intransitive verbs, giving information such as colour and shape. For example, from Tupinambá: itá-piráo-a 'red rock (rock-red-NOMINAL.SUFFIX)'. Nouns may also occur with other nouns, the second indicating some distinguishing feature of the former. For example, Guajajára: ka71-a 7fr 'young monkey (monkey-child)'; Tupinambá ará-ákál 'macaw distinguished by its head (macaw-head)'; Wayampi: pékó-ákká-mirá 'red-headed woodpecker (woodpecker-head-red)'; Parintintin: -ata-tíŋ 'smoke (fire-white)'.

7.4 Indication of number

Tupí-Guarani languages do not have a common plural morpheme. There are some suffixes which are used to indicate a group. For example: *-pér35 (occurring with humans), *-týr and perhaps *-týf. The word *ybytýr 'hill' is a derivation of *yby 'dirt' + *týr. Some languages have their own plural morphemes: Wayampi kó and Emerillon kom were borrowed from a Carib language. In Guajajára a clause-final clitic wa is used when third person subject, object or both are plural and animate (Harrison 1986: 431); this morpheme has a cognate wá in Kayabi.

Another strategy for indicating number is the use of a quantifier, as in (46), although the numerical system in proto-TG was rudimentary. Another strategy is
the use of the diffuse suffix, as in Tupinambá kaʔá-flo ‘throughout the jungle’ in (37), when the subject is countable.

### 7.5 Noun phrases

Noun phrases tend to be short and consist of three basic types: specifier–noun, genitive–noun and noun–appositive,

as in examples (46) and (47), (48), and (49), respectively, from Guajajára (Bendor-Samuel 1972: 110–11). The specifier may be a quantifier or a demonstrative (note that example (47) has two specifiers). In the genitive construction, the r- prefix occurs when the possessed stem is class II. These phrase types may also combine, as in example (49).

(46) mokóz kaʔáʔýr
    two monkey-child
    ‘two young monkeys’

(47) kwéz omó arár
    that certain macaw
    ‘that particular macaw (parrot)’

(48) mamáź r-emáw
    mother LINK-pet
    ‘mother’s pet’

(49) [mamáź r-emáw] zapukáź
    mother LINK-pet chicken
    ‘mother’s pet chicken’

### 7.6 Possession classes

Some stems are grammatically unpossessible, reflecting the indigenous world view, such as the sun, the moon, the jungle. Others are grammatically unpossessible, though they may in fact be possessed in real life, such as ‘chicken’ in example (49).

In this case possession is handled through apposition.

Still other stems are obligatorily possessed, regardless of whether the possessor is in focus. These include body parts and personal possessions. When the possessor is not in focus, an unspecified possessor morpheme is still required. Stems are further subclassified on the basis of the different forms that this morpheme takes, as in table 5.6.

Examples (50)–(53) are from Tupinambá:

(50) (la) i-akáŋ ‘his/her head’ akáŋ ‘(possessed) head’
(51) (lb) i-poʔýr ‘his/her beads’ moʔýr ‘(possessed) beads’
(52) (lIa) s-esá ‘his/her eye’ t-esá ‘(possessed) eye’
(53) (IIId) s-apé ‘his/her path’ pé ‘(possessed) path’

In some languages, like Mbyá, the strategy in example (53) has been replaced by the one used in (52), resulting in forms like t-apé ‘path’.

### 8 Postpositions

Postpositions are similar in form to genitive constructions. Their objects are indicated by the same person markers that are used to indicate the genitive. This is normally by a person marker from set 2 (54 and 56) or a noun (55 and 57).

(54) *i-pypé

<table>
<thead>
<tr>
<th>Stem class</th>
<th>Third person</th>
<th>Unspecified possessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>*ʔi-</td>
<td>*ʔo-</td>
</tr>
<tr>
<td>lb (p-initial)</td>
<td>*ʔi-</td>
<td>*nasalization of initial C</td>
</tr>
<tr>
<td>Ia</td>
<td>*ʔts-</td>
<td>*ʔt-</td>
</tr>
<tr>
<td>lb</td>
<td>*ʔt-</td>
<td>*ʔt-</td>
</tr>
<tr>
<td>Ile</td>
<td>*ʔts-</td>
<td>*ʔt-</td>
</tr>
<tr>
<td>lld (V-initial)</td>
<td>*ʔts-</td>
<td>*subtraction of initial V</td>
</tr>
</tbody>
</table>

Further subclassified on the basis of the different forms that this morpheme takes, as in table 5.6.

15 Some languages have pupé and others pypé. The latter allows for the possible interpretation of this morpheme as a derivation of pyʔa-pe (liver-at).
There are two basic types of person-marking systems in Tupi-Guarani verbs. In independent verbs there is an active-stative marking system for intransitive verbs and a mixed system governed by a person hierarchy for transitive verbs. In other verbal constructions there is an ergative-absolutive system, which also includes the use of coreferential prefixes when appropriate. In some languages the use of the former system has been extended to include other constructions (Jensen 1990).

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.

9.1 Independent verbs
Active intransitive verbs receive A/S prefixes from set I (60). Inactive verbs receive O/S prefixes from set 2, as in examples (61) (class I) and (62) (class II).

In transitive verbs there is a person hierarchy (1 > 2 > 3) which indicates whether the verb should be cross-referenced with A, O, or a combination of the two. Hierarchically superior O is marked with person markers from set 2, as in examples (63) (class I) and (64) (class II). When O is third person, the verb receives an A prefix from set 1, followed by a third person prefix from set 2, as in examples (65) (class I) and (66) (class II). When A is first person and O is second person, special prefixes from set 4 are used which distinguish the number of O.
9.2 Oblique-topicalized verbs

The oblique-topicalized (OBL.TOP) verb construction is used when some sort of setting (time, location, or other circumstance) occurs in the initial position of the sentence as an adverb (68), postpositional phrase (69) or subordinate clause (70). In this case the OBL.TOP verb occurs at the end of the (independent) clause, receives absolute markers from set 2, and receives a special OBL.TOP suffix (-i with consonant-final stems and -i, -w or -j, depending on the language, with vowel-final stems). In some languages, such as Tupinambá and Kayabí, this construction occurs only when A or S is third person. In others, such as Guajajára and Asurini of Tocantins, it occurs when A or S is first or third person, but never when it is second person. In some languages only a few remnants of this system are retained.

(68) (Tb) [kwesé] pajé sjé suşán-i
Yesterday shaman lsg suck-OBL.TOP
'Yesterday the shaman sucked (to cure) me.'

(69) (A) [Tokoróhi(-a) -pe] i-qa-i
Tucurúi to 3-go-OBL.TOP
'He/she went to Tucurúi.'

(70) (A) [ene i-nópo-ramó] i-qa-potar-i
2sg 3-hit-if 3-go-FUT-OBL.TOP
'If you hit him, he will go away.'

38 For a fuller discussion, see Jensen (1987 and 1998b).

9.3 Serial verb constructions

When two or more actions are perceived as part of a whole and the subject of both is the same, they generally occur in a serial verb construction. If the dependent serial verb is transitive (71) or active intransitive (72), it receives a suffix having three allomorphs: *-jiõo (with vowel-final stems), *-a (with consonant-final stems) and *-ta (with stems ending in a diphthong EY). If it is an inactive intransitive verb (73), it receives a different suffix, with allomorphs *-ramo and *-amo. The transitive verb combines with person markers from set 2, referring to O, or with a noun. The intransitive dependent verb receives coreferential person markers since its S is the same as that of the independent verb. Leite (1987) cites the following examples from Tapirapé:

(71) wyrãi ara-pyýk i-fokã-wo i-qi-wo
bird lexcl-catch 3-kill-SER.V 3-eat-SER.V
'We caught the bird, killed it and ate it.'

(72) a-jaok we-yytãp-a
1sg-bathe 1sg.COREF-swim-SER.V
'I'll bathe and I'll swim.'

(73) fe-kaneé i-tyá-ramó
Isg-tired 1sg.COREF-hungry-SER.V
'I am tired and hungry.'

In some languages the coreferential prefixes have been replaced by prefixes from set 1 (Jensen 1990) and the serial verb suffix has been deleted. For example, a-jivy a-jo
'I returned, coming' from the Amapari dialect of Wayampi.

9.4 Temporal subordinate clause constructions

Verbs of temporal subordinate clauses are marked according to an ergative-absolutive system (by set 2 markers or by a noun immediately preceding the verb) and occur at the end of the clause, followed by the subordinating morpheme *-iře ~ *-(ri)re 'after' or *-Vmv ~ -rVmv 'if, when'.

39 What I am calling a dependent serial verb has been referred to as a dependent verb and also (in Portuguese) a gerund. In Portuguese this term refers to what in English would be a present participle, with an adjectival rather than nominal function. In traditional English grammar this term implies that the construction occurs as a noun in contrast to a present participle. According to English terminology, I do not see the term 'gerund' to be appropriate, although there is some similarity to certain nominalizing suffixes (-a, -ta) and to the attributive suffix (-ramo, -amo).
(74) (A) i-ha-ramo oro-karo-póta
   3-go-when 1excl-eat-FUT
   ‘When he goes, we will eat.’

(75) (A) se-nópo-ramo ere-póka
   1sg-hit-when 2sg-laugh
   ‘When he hit me (with an instrument), you laughed.’

If the referent of the person marker (i.e. S, S₁ or O) is identical with the subject of the independent clause, coreferential markers may be used, depending on the extent to which these are used in a particular language.

(76) (A) we-nópo-ramo a-ha-póta (ise)
   1sg.COREF-hit-if 1sg-go-FUT 1sg
   ‘If he hits me (with an instrument), I will go away.’

In some languages the person-marking on subordinate verbs has become like that of independent verbs, as in Mbyá:

(77) a-poranú ramo o-moβá
   1sg-ask when 3-answer
   ‘When I asked he answered.’

9.5 Valency-changing devices

9.5.1 Causatives

The Tupí-Guarani languages have three types of causatives: *mo-, which combines with intransitive verbs, nouns and even nominal suffixes, resulting in a transitive verb (78); *(e)ro-, a comitative causative, which combines with active intransitive verbs, resulting in a transitive verb (79); and *-ukár, which combines with transitive verbs, resulting in a ditransitive verb (80).

(78)(P, Mb) mo-ɾé ‘CAUS-say’ ‘teach’
   (Mb) mo-akú ‘CAUS-hot’ ‘(to) heat’
   (P, Kb) mo-ɾí ‘CAUS-small’ ‘(to) pound, crush’

(79) (W) ero-ɾá ‘CAUS-fall’ ‘to fall and cause (O) to fall’

(80) (W) juká-uka ‘kill-CAUS’ ‘to cause to kill (O)’

In (80) the original O continues to be O and the original A becomes an indirect object.

9.5.2 Detransitivizers

Two morphemes, *je- ‘reflective’ and *jo- ‘reciprocal’, may occur with transitive verbs, resulting in an intransitive verb, as in the following examples from Asurini of Tocantins:

(81) o-se-nópo  ‘He hits himself (with an instrument).’
    o-so-nópo  ‘They hit each other.’

9.5.3 Object incorporation

A direct object may be incorporated in a transitive verb construction. If this object does not require a possessor, as in example (82), the verb is detransitivized. If it requires a possessor, the possessor becomes the direct object, as in example (83).

High saliency of an object precludes its incorporation.

(82) (M) a-y-ɾá
    1sg-water-eat
    ‘I drink-water.’

(83) (M) a-i-po-kʰá
    1sg-3-hand-tie
    ‘I tied his hands.’ (lit. ‘I hand-tied him’)

The incorporation of the generic morphemes *poro ‘people’ or *maʔe ‘thing’ also results in an active intransitive verb, as in examples (84) and (85) from Chiriguano.

(84) a-poro-móe
    1sg-people-teach
    ‘I teach.’

(85) a-maʔ-e-juká
    1sg-thing-kill
    ‘I hunt.’

10 Nominalizations

Adverbs and postpositional phrases are nominalized by a reflex of *-tswár or *-nwár.

(86) (Kb) kope-wát ‘the here ones’
    (WJ) aŋeʔe-wa ‘the now ones’
    (Kb) ytu pe-wát ‘the one at the waterfall’

40 In the first case the subject A becomes Sᵃ, in the second case it remains A and the possessor is raised to O.
Verbs are nominalized to indicate the agent (reflex of *-tsár ~ -ár ~ -tár), circumstance (*-tsáP ~ -áP ~ -táP), and, in the case of transitive verbs, object (*emi-) or patient (*-pýr).

(87) (WA) e-mo7é-ar
1sg-teach-AG.NMLZR 'my teacher'

(88) (G) j-uxin-haw
3-kill-CIRCUMSTANCE 'his being killed'

(89) (Gu) j-ni-emí-ápo
linel-LINK-0.NMLZR-make 'our workmanship'

(90) (Kb) i-juka-pýt
3-kill-PATIENT.NMLZR 'the killed one'

The relativizer *'pa?é (or its variation ma?é) occurs in most languages, though the extent to which it is used varies. In some languages, like Guajajára and Kayabi, it only occurs with intransitive (active and stative) verbs, the referent being S (S₁ and S₂), as in example (91).

(91) (Gj) o-hó ma?é 'the one who went'
(Gj) i-má?éhú ma?é 'the sick one'
(Kb) i-mé ma?é 'the married (husband) one'

In other languages, such as Guarayu, Mbyá, Tupinambá and Wayampi, it occurs with transitive verbs as well, as in the two examples in (92), where the referents are A and O, respectively.

(92) (M) je-ceraá ba7é 'the one that took me'
(M) oro-7é ba7é '(that) which we are eating'

In Wayampi, the indirect object may also be relativized, as in example (93).

(93) (W) a-mé7é i-jupe ma?é
1sg-give 3-to RELZR 'the one to whom I gave it'

11 SEMANTICS

In the Tupi-Guarani languages the centre of the emotions is considered to be the liver (*py7á). A number of constructions, all containing the *py7á morpheme, appear in various languages to communicate feelings, as illustrated by examples (94)–(98):

(94) (Gj) ze-mu-py7á (REFL-CAUS-liver) 'to doubt'

(95) (M) je-mo-py7á (REFL-CAUS-liver) 'to plan, resolve'
py7á-gua7xu (liver-big) 'to be courageous'
py7á-porá (liver-good) 'to be happy'

(96) (P) py7á-mo-týb (liver-CAUS-tighten) 'to be happy, amused'

(97) (U) py7á kaní (liver lose) 'to forget'
py7á ryri7 'to be afraid'
py7á-hu (liver-big) 'to be unafraid'
py7á katú (liver good) 'to love'
py7á monó (liver send) 'to long for'

(98) (WA) ji-py7a-mo7éta (REFL-liver-speak) 'to think'
py7á-kátu (liver-good) 'to be at ease'

It is also possible that the morpheme py which occurs in some words is also derived from *py7á, such as: pyá7á (liver-hard?) 'strong, diligent' in Tupinambá, and pyr7é (liver-at?) 'in, inside of' in several languages.

The concept of pairs (irá 'pair') is basic to the numerical systems of several Tupí-Guarani languages, as can be seen in the examples (99)–(102).

(99) (W) iró '(to be) even'
níroi '(to be) odd' (lit. 'not even')
iróte 'four'
iróíróté 'six or eight'

(100) (A) na7irú7 'three' (lit. 'not even')

(101) (Gj) ná7irú7 'three' (lit. 'not even')

(102) (M) irú 'partner'
je-mo-irú 'to be a partner'
irúny 'four'

BIBLIOGRAPHY

Anchieta, J. de. 1595. Artes de gramática da língua mais usada na costa do Brasil. Coimbra: Antonio de Mariz. (There are various reproductions of this work, including one by the Biblioteca Nacional, Rio de Janeiro, 1933, and one by Editora Anchieta, Sao Paulo, 1946.)

Macro-Jê is the name that was proposed about fifty years ago by Mason (1950: 287) for a large array of South American languages (all spoken in Brazil) thought to be related to the Jê linguistic family. W. Schmidt (1926: 234–8) had earlier employed the name Ges-Tapuya, and Loukotka (1944[1942]: 2–6) the name Tapuya-žê in the same sense. The possible genetic relationship among the many languages ascribed to the Macro-Jê stock is a working hypothesis whose details have varied according to different scholars. Loukotka included in it the following eight 'language families': žê (Jê), Opie (Ofayé), Kaingân (Kaingáng), Coroado (Purí), Mašakali (Maxakalí), Patašô (Pataxó), Botokudo (Krenák) and Kamakan (Kanaká) (names or spellings added in parentheses are those now currently used by Brazilian linguists and anthropologists and adopted in the present chapter). Nimuendajú (Métraux and Nimuendajú 1946: 542; Nimuendajú 1945[1980]) considered Malalí an independent linguistic family, whereas Mason (1950) added Malalí and Koropó (Koropó) to Macro-Jê, but took out Ofayé as well as Ia té (Yatê). The latter had already been separated from the stock by Loukotka (1942). Later, on the basis of good comparative work, Davis (1966) demonstrated that Kaingáng is truly a member of the Jê genetic family and not another family in itself. In another paper, Davis (1968) gave evidence of regular phonological correspondences between Jê and Maxakalí as well as between Jê and Karajá and mentioned possible wider relationships of Macro-Jê with Boróro, Tupí, and Fulnió (Yatê). As for Boróro, Guérios (1939) presented as indicative of genetic relationship the similarities he had found between Eastern Boróro and two Northern Jê languages, Timbira ('Merrime' = Canela) and Kayapó. Gudschinsky (1971), comparing Ofayé with Davis' reconstruction of proto-Jê, showed that it is more likely to be a member of Macro-Jê. Boswood (1973) gave some lexical evidence in favour of the inclusion of Rikbaktsá in this stock. Rodrigues (1986) included in it Kariri and Guató, but dealt with Pataxó as a member of the Maxakalí family, the same being true of Malalí, which had already been put in this family (and not as a main branch of Macro-Jê) by Loukotka. For Greenberg (1987), all the aforementioned languages or language families (except for Kariri) belong to Macro-Jê, as do
Chiquito, Otí and Yabutí (Jabutí). Although Greenberg (1987: 86) states that 'these three languages are as validly Macro-Ge as the others', the meagre data he presents for Otí and Yabutí do not substantiate his claim. Kaufman, in his recent revision of the classification of South American languages (1990, 1994), left out of Macro-Jé these two languages as well as Kariri, but retained Chiquito. In the present chapter Kariri, but not Chiquito, is considered a possible member of Macro-Jé.

A part of the languages involved in the Macro-Jé hypothesis are already dead and most of them have been very poorly documented. Thus it is very difficult to work out their relationships with the best-known members of the stock. Table 6.1 lists the language families and the single languages that will be considered here as possible members of Macro-Jé (for which there is some documentation, albeit very scarce in some cases). For some selected lexical and phonological evidence of the consistency of the whole group see §7 below. Table 6.1 also gives information on the locations of the languages and the number of speakers. Names of dead languages are preceded by X. The approximate locations of the extant Macro-Jé languages are shown on map S.

2 DISTRIBUTION

Although several Macro-Jé languages are spoken in Brazilian Amazonia, the geographical distribution of this linguistic stock is rather circum-Amazonian, encircling Amazonia on its eastern and southern sides. Most members of the Macro-Jé linguistic stock have been spoken in eastern and northeastern Brazil, but a few language families are found in central and southwestern Brazil. With the exception of Otuque of the Bororo family, spoken west of the Paraguay river in Bolivia, the whole Macro-Jé stock is found entirely in what is today Brazilian territory. In the sixteenth century, when the Europeans (mainly Portuguese and French) started trading or settling on the Atlantic coast of South America, most of this coast was occupied by Indians speaking languages of the Tupí-Guaraní family. Some places people speaking Macro-Jé languages used to go to the seashore seasonally in order to gather seafood. This was probably the case with the Maromomim or Guarulho on the coast of present-day São Paulo State and with the Waitaká on the coast of northern Rio de Janeiro and Espirito Santo, but it is likely that many other peoples would have done the same. Both the Maromomim and Waitaká spoke languages that probably belonged to the Puri family of Macro-Jé, but both languages became extinct and no record of them has been preserved or, at least, has not so far been found. Of the language of the Maromomim there is clear historical information that a grammar and a dictionary as well as a Christian doctrine, were written in the sixteenth century by the Jesuit Manuel Viegas in cooperation with Joseph de Anchiesta, the linguistically skilled author of the first grammar of the Tupí-Guaraní language Tupinambá.

From a geographical point of view the Macro-Jé languages may be divided into

6 Macro-Jé

<table>
<thead>
<tr>
<th>Table 6.1 Macro-Jé languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>I  Je family</td>
</tr>
<tr>
<td>a North-eastern Je</td>
</tr>
<tr>
<td>1 Taíko (south-eastern Piauí)</td>
</tr>
<tr>
<td>b Northern Je</td>
</tr>
<tr>
<td>1 Timbira (including Canela Ramkokamekrá, Canela Apanyekrá, Gavião Piokójé, Gavião Parakatejé, Krinkati, Krabó, Krené; Maranhão, Pará, Tocantins; 2,800)</td>
</tr>
<tr>
<td>2 Apinajé (northern Tocantins; 720)</td>
</tr>
<tr>
<td>3 Kayapó (including Aukré, Gororité, Karanáo, Kikretem, Kokraimóéro, Kukrenkранé, Menkranugnoi, Mentuktire, Xikrin; eastern Mato Grosso, south-eastern Pará; 5,000)</td>
</tr>
<tr>
<td>4 Panará (Kren-akájóvé) ( Área Indígena Panará, northern Mato Grosso and south-western Pará; 160)</td>
</tr>
<tr>
<td>5 Suyá (including Tapayuna; Xingu Indigenous Park in Mato Grosso; 213 S., 58 T)</td>
</tr>
<tr>
<td>c Central Je</td>
</tr>
<tr>
<td>1 Xavante (south-eastern Mato Grosso, formerly western and northern Goiás; 9,000)</td>
</tr>
<tr>
<td>2 Xerénte (Tocantins; 1,550)</td>
</tr>
<tr>
<td>3 ? Xikriába (Minas Gerais; 7,700 ethnic, probably no speakers)</td>
</tr>
<tr>
<td>4 ? Akró (eastern Goiás, southern Maranhão)</td>
</tr>
<tr>
<td>d Southern Je</td>
</tr>
<tr>
<td>1 Kaingáng (including São Paulo K., Paraná K., Central K., South-western K., and South-eastern K.; São Paulo, Paraná, Santa Catarina, Rio Grande do Sul; 20,000)</td>
</tr>
<tr>
<td>2 Xokléneg (Santa Catarina; 1,650)</td>
</tr>
<tr>
<td>3 ? Ingaín (north-eastern Argentina, south-eastern Paraguay)</td>
</tr>
<tr>
<td>e Kamaká family</td>
</tr>
<tr>
<td>1 ? Kamaká (south-eastern Bahia, north-eastern Espirito Santo)</td>
</tr>
<tr>
<td>2 ? Mongoyó (south-eastern Bahia)</td>
</tr>
<tr>
<td>3 ? Menién (south-eastern Bahia)</td>
</tr>
<tr>
<td>4 ? Koxoxó (south-eastern Bahia)</td>
</tr>
<tr>
<td>5 ? Masakkará (north-eastern Bahia)</td>
</tr>
<tr>
<td>III Maxakali family</td>
</tr>
<tr>
<td>1 Maxakali (north-eastern Minas Gerais, northern Espirito Santo; 854)</td>
</tr>
<tr>
<td>2 ? Kapoxó (including Kumanaxó and Paháme; north-eastern Minas Gerais and south-eastern Bahia)</td>
</tr>
<tr>
<td>3 ? Monoxó (north-eastern Minas Gerais and south-eastern Bahia)</td>
</tr>
<tr>
<td>4 ? Makó (north-eastern Minas Gerais)</td>
</tr>
<tr>
<td>5 ? Malakó (north-eastern Minas Gerais)</td>
</tr>
<tr>
<td>6 ? Pataxó (including Hähaháe; south-eastern Bahia; 4,600 ethnic)</td>
</tr>
<tr>
<td>IV Krenák family</td>
</tr>
<tr>
<td>1 Krenák (formerly called Botocudo, including Nakrehé, Nakpié, Nakyanák, Nakyanapá, Neypneyp, Etwet, Mínuyiún, Yiporók, Pójító, Poten, Kreekmún, Bakue, Araná; north-eastern Minas Gerais and northern and central Espirito Santo, formerly also south-eastern Bahia; some families now also in central São Paulo; 100 ethnic, about 10 speakers)</td>
</tr>
<tr>
<td>2 ? Guéren (south-eastern Bahia)</td>
</tr>
<tr>
<td>V Puri family</td>
</tr>
<tr>
<td>1 Puri (Espirito Santo, Rio de Janeiro, north-eastern São Paulo, south-eastern Minas Gerais)</td>
</tr>
</tbody>
</table>
Table 6.1 (cont.)

<table>
<thead>
<tr>
<th>Macro-Jê Family</th>
<th>Yate</th>
<th>Kamaká</th>
<th>Maxakalí</th>
<th>Rikbaktsá</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Kamaká family</td>
<td>1 Kamaká (between the Jequitinhonha and Contas rivers in south-eastern Bahia)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II yate family</td>
<td>1 Yaté (the people are named Fulnió, formerly known as Carniú; Pernambuco; 3,000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III Karirí family</td>
<td>1 Karirí (including Southern K., Northern K., Javari and Xambioá; eastern Mato Grosso; western Goyaz; 3,000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV Bororo family</td>
<td>1 Eastern Bororo (western Mato Grosso; 1,072 ethnic, some of which no longer speak Bororo)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V Xikaray family</td>
<td>2 Western Xikaray (eastern Mato Grosso)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI Yate (the people are named Fulnió, formerly known as Carniú; Pernambuco; 3,000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII Kamaká (between the Jequitinhonha and Contas rivers in south-eastern Bahia)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIII Karirí (including Southern K., Northern K., Javari and Xambioá; eastern Mato Grosso; western Goyaz; 3,000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX Bororo family</td>
<td>1 Eastern Bororo (western Mato Grosso; 1,072 ethnic, some of which no longer speak Bororo)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Xikaray family</td>
<td>2 Western Xikaray (eastern Mato Grosso)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XI Guató</td>
<td>1 Guató (south-western Mato Grosso; 380 ethnic, only about 5 speakers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XII Rikbaktsá family</td>
<td>1 Rikbaktsá (northern Mato Grosso; 900)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Eastern, central and western, extending diagonally across the South American lowlands from the eastern and north-eastern Atlantic coast to the upper Paraguay river.

The central group comprises the Karajá and Jê families, the latter spreading northwards and southwards along the savanna (cerrados and campos) plains to Piaui and Maranhão in the north and to Rio Grande do Sul in the south. Karajá extends also on a north-south axis, but is restricted to the valley of the Araguaia river, including Bananal island.

Eastern Macro-Jê languages are those belonging to the Puri, Krenak, Maxakali, Kamaká, Kariri and Yaté families. The western families are Ofayé, Bororo, Rikbaktsá and Guató. The distribution of the eastern Macro-Jê language families extending from eastern to north-eastern Brazil was as follows (taking the modern Brazilian states and some main rivers as geographical references). Puri was along the whole extent of the Paraíba do Sul river and spreading northwards until the Doce river, in north-eastern São Paulo, Rio de Janeiro, south-eastern Minas Gerais, and southern Espirito Santo. Krenak extended from the Doce river to the Paraguacu river in north-eastern Minas Gerais, central and northern Espirito Santo, and south-eastern Bahia. Maxakali went from the Sapucai and São Mateus rivers to the Jequitinhonha river in eastern Minas Gerais, northern Espirito Santo and south-eastern Bahia. Kamaká was between the Jequitinhonha and Contas rivers in south-eastern Bahia. Kariri was located mainly between the Itapicuru and the middle and lower São Francisco river, in central and north-eastern Bahia and southern Sergipe, and with some extensions northwards and southwards, probably due to the Portuguese invasion of their territory in the seventeenth century. Yaté was north of the lower São Francisco, in eastern Pernambuco.

Of the western families, Bororo spread from the sources of the Araguaia river into eastern Bolivia, and Guató was found on the upper Paraguay, immediately to the south-west of Bororo. Ofayé occupied a stretch of the upper Paraná river from the mouth of the Tietê downwards, passing the mouth of the Paranapanema and extended to the west to the headwaters of the right-bank tributaries of the Paraná, reaching also the headwaters of some tributaries of the Paraguay river. Rikbaktsá is located far to the north, on the upper Juruena river, one of the headwaters of the Tapajós. Rikbaktsá is the only language of the Macro-Jê linguistic family to be found exclusively in Amazonia. However, some peoples speaking languages of the Jê family also live in Amazonia, such as the Panará, the Suyá, the Tapayuna and the Kayapó. It is likely that all of the latter entered Amazonia coming from the east under the pressure of the Portuguese invasion of the savannas in central Brazil. In the eighteenth century the Panará mainly lived around the headwaters of the Paraná river, such as the Paraíba and the Grande, in present-day southern Goiás, south-western Minas Gerais and north-western São Paulo, where they came to be known as Kayapó. After their disappearance due to the war waged against them by the Portuguese, this name was applied to another Jê people, whose language is similar to, but not the same as, that of the Panará. Very recently (1968) a large population of Panará was discovered on the Peixoto de Azevedo river, another headwater of the Tapajós, where they had taken refuge and managed to live for more than two centuries in freedom and peace. Shortly after having been forced into a new contact with white people in the early 1970s, the Panará population was drastically reduced by epidemics. The survivors were transported to the Xingu Indigenous Park where they have since lived, but they are now managing to get back to a part of their former territory. The modern Kayapó moved into eastern Amazonia during the first half of the twentieth century, but they were preceded in this move by the Suyá, who were already on the Upper Xingu in the second half of the nineteenth century.

3 Linguistic Scholarship

A significant number of Macro-Jê languages became extinct as a consequence of the European settlement in Brazil. The worst-hit language families were those located in...
eastern Brazil, since they were most affected by the Portuguese gold-seeking and slave-hunting expeditions. Although the very first victims of the European expansion were the Tupi-Guarani-speaking Tupinambá, who during the sixteenth century were wiped out on large stretches of the Atlantic coast, some Macro-Jê-speaking peoples became the objects of Portuguese attacks around this time — for example, the Aimoré in southeastern Bahia (whose language probably belonged to the Krenak family). At the beginning of the nineteenth century there were still small groups of Indians in eastern Brazil who spoke languages of all the families here ascribed to Macro-Jê, but many other languages were already extinct and we do not know whether they were indeed Macro-Jê or not. The first years of the nineteenth century were characterized by the beginning of the study of Brazilian natural history by European scientists, mainly Germans and French. Many botanists, zoologists and geologists crossed the country gathering samples of the flora, the fauna and the soils and also gathered samples of the languages spoken by the Indians they met on their way. At that time there were no specialists in the study of languages and the model of language documentation was the collection of short comparative lists of words, such as those published in the eighteenth century by Adelung and by Hervás. The scientific curiosity of the naturalists has saved from complete oblivion scores of words of many languages that, during the following years, became extinct.

All the languages of the Puri, Kamaká and Kariri families are now dead. Two languages of the Kariri family are quite well known thanks to the work of seventeenth-century missionaries. An Italian Jesuit published a fine grammar and a catechism in Kipeá (Mamiani 1698, 1699) and a French Capuchin published a catechism in Dzubukú (Bernardo de Nantes 1709). Knowledge of the Puri language is limited to the scanty data in lists gathered by European naturalists and by a Brazilian engineer (for the linguistic data see Loukotka 1937). The Kamaká language was the last of the family with the same name to disappear; it was still possible to elicit words and phrases from the last speakers in the 1930s and 1940s (Guérios 1945, 1948; for earlier data see Loukotka 1932). For the Krenak family there are today no more than ten speakers representing two or three dialects of the one language (information on published and unpublished sources in Seki 1990; see also Silva 1986 and Sebestyén 1981). Some of these, speaking the Nakrehé dialect, live far from their original home, in the interior of the State of São Paulo, to where they were deported by the Brazilian administration about forty years ago. The Maxakalí language is the only surviving member of the Maxakalí family, the other five languages now being dead (recent studies are Pereira 1992; Popovich 1967, 1971, 1985, 1986; Gudschinsky, Popovich and Popovich 1970; Rodrigues 1981; Wetzel 1996; Wetzels and Sluyters 1996; for earlier data see Loukotka 1931 and 1939). Yaté is the only indigenous language to survive in northeastern Brazil, in a region where there are remnants of several other indigenous peoples who now speak only Portuguese. It is not known whether the extinct languages of some of these peoples belonged to the same family as Yaté or even whether they were members of the Macro-Jê stock (studies on Yaté are Lapenda 1968; Pinto 1956: 265–76 and passim; Meland 1968; Meland and Meland 1967, 1968; Barbosa 1991).

The languages of the Jê family were better preserved due to the inland location of their speakers. Apparently only one main branch of the family disappeared completely, namely Jaikó, whose living area in the backlands of Piauí was one of the first to be used by the Portuguese for intensive cattle-raising. The other three subgroups still have some healthy languages, despite suffering heavy population losses. For Jaikó the only source is a small word list gathered and published by Von Martius (1867: 143); for northern Jê the main contributions are Popjes and Popjes (1986), Souza (1989) (Timbirá); Callow (1962), Ham (1961, 1967), Ham, Waller and Koopman (1979) (Apinajé); Stout and Thomson (1974a, b), Thomson and Stout (1974), Jefferson (1989), M. de N. de O. Borges (1995) (Kayapó); Dourado (1990, 1993a, b) (Panará); Guedes (1993), Santos (1997) (Suyá); for central Jê: Mattos (1973) (Xerente); Hall (1979), Hall, McLeod and Mitchell (1987), McLeod (1974), McLeod and Mitchell (1977, 1988) (Xavante); for southern Jê: Guérios (1942), Mullern (1965, 1966), Wiesemann (1971, 1972, 1978), Rodrigues and Cavalcante (1982), Cavalcante (1987), Teixeira (1988) (Kaingang); Henry (1935, 1948), Urban (1985), Bublitz (1994) (Xokleng).

The main literature on Karajá comprises Ehrenreich (1949), Kunike (1916, 1919), Fortune (1973), Fortune and Fortune (1975), Maia (1986), Cavalcante (1992), Ribeiro (1996), Borges (1997). The only analytical essay on Ofayé is Gudschinsky (1974), which makes reference to previous sources. For the Bororo family there is a grammatical essay and a monumental encyclopaedic dictionary plus the text collections by the members of the Salesian Mission (Colbacchini 1925; Colbacchini and Albisetti 1942, Albisetti and Venturelli 1962, 1969, 1976), as well as Crowell’s grammar (1979), all for Eastern Bororo. For Umutína see M. Schmidt (1941), Schultz (1952) and Lima (1993); Rodrigues (1962) provides a comparison of Umutína with Eastern Bororo. Guató was analysed by Palácio (1984, 1986), who gives information on previous data. For Rikbaktsá see Boswood (1974a, b, 1978) and Tremaine (1981).

4 PHONOLOGY

4.1 Vocalic systems

As in other language families of Lowland South America, a very common feature in many languages of the Macro-Jê stock is the presence of phonologically contrastive nasal vowels. Often it is the nasal vowels that condition the variation of the
consonants and not the reverse (as happens in other languages). In general the number of nasal vowels is less than the number of oral ones. The Paraná dialect of Kaingang (Jê), for instance, has nine oral and five nasal vowels, as may be seen in table 6.2. In this language the low central nasal vowel oscillates from rounded back [o] to unrounded central [a]. All the languages of the Jê family have vowel systems as large as that of Kaingang or with one or two more vowels. Apinajé, which distinguishes four instead of three degrees of height for central unrounded, has ten oral and seven nasal vowels as shown in table 6.3.

The system of nine or ten oral vowels and a lesser number of nasal vowels is typical of the languages of the Jê family; Davis (1966) reconstructed a system of nine oral and six nasal vowels for proto-Jê. A system comparable in the number of oral vowels is found in the Makú family (see chapter 9). The languages of the other families of Macro-Jê vary between nine and five oral vowels, but in general preserve the distinction of three tongue positions (front, central and back). Some of them have contrasting length. Karajá has nine oral vowels like Kaingang, but has only two nasal vowels, ş and ô. Phonetically there is also [i], which is an automatic realization of the phoneme /l/ when it either stands at the beginning of a word or is preceded by /b/ or by a voiced stop, as in the following examples: adil [adil] 'your mother', adil [adil] 'grass', habu [habu] 'man', su [șə] 'armadillo', bahadu [məhədu] 'group', madi [mədil] 'my mother', šada [șədə] 'face'.

Kipeá, of the Kariri family, has seven oral and five nasal vowels, as shown in table 6.4 (after Azevedo 1965), whereas Maxakali has only five in each category, as may be seen in table 6.5 (after Gudschinsky, Popovich and Popovich 1970).

According to Gudschinsky (1974), Ofayé has seven oral and four nasal short vowels and as many long vowels (see table 6.6, after Gudschinsky 1974), Guató has eight oral and five nasal (table 6.7, after Palácio 1984), and Rikbaktsá has six oral and six nasal vowels (table 6.8, after Boswood 1973). Yaté and Boróro have no nasal vowels and have only seven oral vowels each, but

| Table 6.2 Vowel system of the Paraná dialect of Kaingang (after Wiesemann 1972) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | oral            | nasal           |                 |                 |                 |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |

| Table 6.3 Vowel system of Apinajé (after Ham 1967) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | oral            | nasal           |                 |                 |                 |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |

| Table 6.4 Vowel system of Kipeá |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | oral            | nasal           |                 |                 |                 |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |

| Table 6.5 Vowel system of Maxakali |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | oral            | nasal           |                 |                 |                 |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |

| Table 6.6 Vowel system of Ofayé |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | short           | long            |                 | short           | long            |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |
| Table 6.7 Vowel system of Guató |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | short           | long            |                 | short           | long            |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |
| Table 6.8 Vowel system of Rikbaktsá |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | short           | long            |                 | short           | long            |
|                 | front           | central         | back            | front           | central         | back            |
|                 | unrounded       | unrounded       | rounded         | unrounded       | unrounded       | rounded         |
| i               | i               | ü               | ü               | ü               | ü               |
| e               | a               | o               | ü               | ü               | ü               |
| ê               | a               | o               | ê               | ü               | ü               |


Kipeá, of the Kariri family, has seven oral and five nasal vowels, as shown in table 6.4 (after Azevedo 1965), whereas Maxakali has only five in each category, as may be seen in table 6.5 (after Gudschinsky, Popovich and Popovich 1970).

According to Gudschinsky (1974), Ofayé has seven oral and four nasal short vowels and as many long vowels (see table 6.6, after Gudschinsky 1974), Guató has eight oral and five nasal (table 6.7, after Palácio 1984), and Rikbaktsá has six oral and six nasal vowels (table 6.8, after Boswood 1973). Yaté and Boróro have no nasal vowels and have only seven oral vowels each, but
Table 6.7 Vowel system of Guató

<table>
<thead>
<tr>
<th></th>
<th>oral</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>central</td>
<td>back</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
<td>u</td>
</tr>
</tbody>
</table>

Table 6.8 Vowel system of Rikbaktsá

<table>
<thead>
<tr>
<th></th>
<th>oral</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>central</td>
<td>back</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 6.9 Vowel system of Yaté

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>central</td>
</tr>
<tr>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>e</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 6.10 Vowel system of Boróro

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>central</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>e</td>
<td>a</td>
</tr>
</tbody>
</table>

both partially preserve the three-column distinction as well as the three-row distinction of the fuller Macro-Jê systems, as may be seen in table 6.9 (after Meland and Meland 1967) and table 6.10 (after Crowell 1979).

4.2 Consonantal systems

The consonantal systems of Macro-Jê languages are of medium size. Only Yaté has slightly more than twenty consonants incorporating a full set of aspirated and unaspirated voiceless stops. Besides these it has two voiced stops and five fricatives, two nasals and three approximants, as shown in table 6.11. Kipeá (Kariri family), Yaté's neighbour to the south, had no aspirated stops, but had a full series of voiced stops and also a glottal stop, only two fricatives, but four nasals and three approximants. Note that Yaté has a lateral approximant, whereas Kipeá has a central flap, like most Macro-Jê languages (see table 6.12).

The consonantal system of Guató resembles those of Yaté and Kipeá, but has no series of aspirated stops and no alveolar stops or fricatives; it adds to its inventory two labio-velar stops, as may be seen in table 6.13. Like Kipeá and Guató, Boróro
Table 6.14 Consonantal system of Bororo (after Crowell 1979)

<table>
<thead>
<tr>
<th>部位</th>
<th>labial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>lt</td>
<td>tf</td>
<td>k</td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>d3</td>
<td>g</td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>w</td>
<td>r</td>
<td>j</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.15 Consonantal system of Maxakalí (based on Gudschinsky, Popovich and Popovich 1970)

<table>
<thead>
<tr>
<th>部位</th>
<th>labial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>lt</td>
<td>tf</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td></td>
<td>g</td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>h</td>
</tr>
</tbody>
</table>

Table 6.16 Consonantal system of Karajá (after Fortune 1973)

<table>
<thead>
<tr>
<th>部位</th>
<th>labial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>tf</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>d3</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td>voiced implosive</td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>approximant</td>
<td>w</td>
<td>r</td>
<td>j</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Maxakalí has a remarkably reduced inventory of ten consonantal phonemes, and also ten vocalic phonemes. It has a full series of voiced and voiceless stops, each in four points of articulation, and two glottal sounds, a stop and a fricative, as may be seen in table 6.15.

The consonantal system of Karajá is also remarkable not only for its reduced inventory of twelve phonemes, but also for its concentration on dental and alveo-palatal sounds and its great dissymmetry. Table 6.16 displays this system.

The voiced stops /b/ and /d/ have fully nasal allophones, [m] and [n], occurring before nasal vowels.

Another phonological peculiarity of Karajá is the systematic differentiation of

Table 6.17 Differences between men's speech and women's speech in Karajá (data from Fortune 1973 and Borges 1997)

<table>
<thead>
<tr>
<th>部位</th>
<th>men's speech</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>voiced implosive</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>w</td>
<td></td>
</tr>
</tbody>
</table>

the phonological shape of words between men's speech and women's speech (Ehrenreich 1894, Krause 1911, Kunike 1916, Fortune and Fortune 1975, Borges 1997). Men's speech regularly lacks the velar stops present in the speech of women, as well as the instances of the voiceless alveo-palatal affricate that are historically derived from velar stops palatalized by a preceding i. As a consequence of the dropping of the velar stops, several vowel contractions may result, thus making the shape of the words and sentences uttered by men more distinct from those pronounced by women. Even borrowings from Portuguese are subject to the dropping of the velar stop. Table 6.17 presents some examples of women's and men's speech in Karajá.

It is not easy to conceive of a situation in which this gender-based systematic differentiation between the phonological shape of the words might have originated. Perhaps in the past the women of one dialectal group of Karajá could have been subjugated by warriors speaking another language, say one lacking velar stops but having glottal stops. Such warriors could have killed all Karajá male adults, taken their place as husbands and learned the Karajá language from their new wives, but substituted their glottal stops for the velar stops of the women. This bad pronunciation by the new masters of the group would then have been maintained through the following generations and spread to other dialectal groups. In the course of time, the articulation of the glottal stop would have weakened and finally disappeared, giving place to
vowel sequences and contractions (as shown in table 6.16, the Karajá language has no glottal stop phoneme). It happens that the neighbouring language Xavante has historically undergone the systematic change of velar consonants into glottal stops (compare Xerente dakar with Xavante da?ra 'somebody's head', Xe. dakkwa with Xa. da?wa 'somebody's tooth', and Xe. ka? with Xa. ?adga 'fire'). This language, which does not show a similar difference correlated with the sex of the speakers, may well be the source of the difference between Karajá men's and women's speech (at least the main difference - the dropping of velar consonants in men's speech).

Table 6.18 shows the consonantal system of Xavante, in which there is no velar phoneme. This is the only Macro-Jé language lacking velars.

Table 6.18  Consonantal phonemes of Xavante (after McLeod 1974)

<table>
<thead>
<tr>
<th></th>
<th>labial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>d</td>
<td>k</td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>d3</td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>w</td>
<td>r</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The consonantal systems of the other languages of the Jé family are in general simple, with a series of four voiceless stops (labial, dental, alveo-palatal and velar) and another of four corresponding nasals, and three approximants; Davis (1966) reconstructed such a system for proto-Jé (the approximants being w, r and z, the latter being an arbitrary symbol for a phonetically undefined proto-phoneme).

Some languages also include a glottal stop and/or a glottal fricative. Timbira has an aspirated velar stop contrasting with the plain one. Kaingáng adds to this system a bilabial fricative, that is voiceless (/>. in most dialects, but is voiced (/β/) in the dialect of São Paulo. Xokleng, which is more similar to Kaingáng, has a voiced interdental fricative corresponding etymologically to the bilabial fricative of Kaingáng, e.g. Kaingáng ñi, Xokleng ñi?iseed'; K ñi, X ñi ñi 'bitter'; K ñi?i, 'to take the breast', X ñiõi, 'to give the breast'; K ño?i, X ñiõi 'skin'. Xokleng also has a lateral approximant where Kaingáng has a central flap, as in the last example and in the following: K ño, X ña 'sun'; K ñi?i, X ñi?i 'white'; K ñi, X ñe?i 'head'. The Jé language with the most complex system is Kayapó (table 6.19), which has a complete series of voiced stops, contrasting with the voiceless ones and the nasals, as may be seen in the following examples: pi 'wood', bi 'to kill', mi 'alligator'.

Some Macro-Jé languages have internally complex phonemes - that is to say, phonemes that are realized by a sequence of sounds (or, in physiological terms, produced by a sequence of articulatory movements). The Southern Jé language Kaingáng is one of the most striking cases. In this language nasal consonantal phonemes are pronounced in up to three phases of articulation, [=nasal] [+nasal] [+nasal] (Rodrigues and Cavalcante 1982, Cavalcante 1987, cf. Anderson 1974, 1976). The word kaner 'smooth' is pronounced [ka?ndere]: between oral vowels the nasal phoneme n starts as an oral stop [d], changes to the nasal [n], and changes back to an oral [d]. The inner phase is fully nasal, whereas the first and third phases are assimilated to the oral vowels bordering the consonant. If one of the vowels is nasal, only the phase contiguous to the oral vowel is oral: kaner 'eye' is pronounced [ka?ndere], whereas eme 'blue sky' is [eme]. If both vowels are nasal, the consonant is fully nasal: eme 'bread' [eme], pani 'back' [pani]. Pause or silence before and after the nasal consonant has the same effect as a nasal vowel: eme 'thing' [eme], nimm 'to give a long object' [nimm], but word boundaries inside an utterance are no obstacle for the assimilation of the first and third phases of the nasal consonants: no [nde] 'arrow', ti no [thin] 'his arrow', no nimm [ndenimm] 'to give arrows'. The first example above, kaner, illustrates another complex phoneme. The phoneme r starts as a vowel at the beginning of a word, as in ra ['ra] 'thin', rä ['rä] 'sun'; and ends as a vowel with the same features as the vowel that precedes it at the word end, as in kar ['kar] 'all', ñor ['ñor] 'full', ñer ['ñer] 'feather, wing'. The approximants w and / show the same behaviour as r, but only in final position: ra?i ['ræi] 'broken', taw ['taw] 'covering', teñ ['teñ] 'long', ra?i ['ræi] 'ripe', wij ['wij] 'bow', jâjara ['jâjara] 'hook'.

Another notorious case of complex phonemes is Maxakalí, whose stop consonants may be realized with fully vocable phonemes and, according to the phonological environment, may actualize only the vocable phase, as in pipkip [pipkip] 'nail', tifik [tfii] 'to cut', batisk jë [mabak jë] 'angry toad', tapet [tapet] 'paper' (p is realized as [e?] in the first example, k as [k] in the second and third, b as [n] in the third, and as [t] in the third and as [tã] in the fourth) (see Gudschinsky, Popovich and Popovich 1970 for the details of the allophonic variation of Maxakalí consonants, Rodrigues 1981 for the nasalization of the consonants, and Wetzel 1996 for an autosegmental interpretation of the data).
Consonantal clusters in most Macro-Jê languages are limited to the combination of grave (labial and velar) stops followed by a central or a lateral approximant, as in Kipeá 'blood', krú 'tail', or in the following Kaingáng examples: pra 'to bite', mrúr 'vine', kru 'rowboat', gref 'to dance'. An exception is Yate, which presents a large variety of sequences of two, three and four consonants - for example: kwlélja 'rotten', ehléndomnja 'star', kénja 'advertising', eþiklkja 'bad odour', eþka 'pulling out', eþikka 'cleaning', etikkja 'blanket', kliæø 'nose', kæliæø 'wind', kæliæø 'message', fækka 'snake', fækka 'knife'.

Another exception is the Central Jê language Xerênte, which has developed very complex consonantal clusters: psedi 'it is good', kdc; 'path', kdc; 'stone', tbe 'fish', kbru 'all', sðikka 'ashes', sð; 'to cut', sra 'hill', zraku 'to the other river bank', tbe 'iron', krdbr5hd5 [krnmr5mn5] 'to live', bø to ikkire 'he got thin' (Mattos 1973).

4.3 Tone

Two Macro-Jê languages have been described as being tonal, Yate in the Brazilian northeast and Guató in the Brazilian southwest. Yate has two distinctive levels of tone, but word-finally the tone tends to be middle; according to Meland and Meland (1967) there is morphophonological 'tone perturbation'. Examples of words contrasting only in tone are (tones marked as follows: á high, a middle, í low): ífia 'wide', ífia 'fast', úsja 'a kind of pigeon', úsja 'centipede', índka 'denouncing', índka 'tasting'; íf 'to suck', íf 'to scratch'. Guató (Palácio 1986) also has only two distinctive levels of tone: mói 'piranha', mói 'word'; mikí 'pot'; mikí 'foot'; mikí 'tobacco', mikí 'dove'. Although none of the best-analysed Jê languages have been reported to be tonal, it is said that Krahó (Timbíra) speakers may communicate at distance by means of 'sentence whistling' (Suell M. de Souza, p. c.).

5 Morphology

Macro-Jê languages are agglutinating and mildly synthetic; they combine elements of head- and of dependent marking.

5.1 Inflection for contiguity of a determiner

The morphology of most Macro-Jê languages is not very complex. Word classes are distinguished more by syntactic strategies than by affixation. A pervasive inflec-
Timbira has no reflexive prefix and has a prefix only for the 1st plural inclusive: pa-
Kipeá and Timbira, and likewise Karajá and Maxakalí, make reference to a
personal possession, e.g. Kaingáng ?¡JI

Sorne other languages use diticized pronouns instead of prefixes for marking per-
sonal possession, e.g. Kaingáng ?¡JI

Some languages, although they have no marker of pluralization on nouns, have
plural pronouns or personal prefixes on the verb for agreement with plural subjects
and some of them also for plural objects. In the Jé family one such language is
Kaingáng (Paraná dialect), which has personal pronouns for 3rd person plural mas-
culine and feminine: 2ap 'they (masc.)', 2ap 'they (fem.)'. Kaingáng also has plural
verbs for agreeing with plural S or O, even though the nouns are not marked for
number. In this language there are several morphological devices for marking plu-
rality on verbs: prefixation, infixation, reduplication, a combination of two of these
with or without ablaut, and also suppletion. See table 6.21 for some examples from
the Paraná dialect.
Table 6.21 Number-marking in Kaingáng  
(Paraná dialect) (from Cavalcante 1987)

<table>
<thead>
<tr>
<th>SG (S or O)</th>
<th>PL (S or O)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 pra</td>
<td>kipra</td>
<td>'to bite'</td>
</tr>
<tr>
<td>2 jànnã</td>
<td>fàngnã</td>
<td>'to use almost all'</td>
</tr>
<tr>
<td>3 ðí</td>
<td>kipfí</td>
<td>'to weave'</td>
</tr>
<tr>
<td>4 kàjãm</td>
<td>kinjãm</td>
<td>'to pay, to buy'</td>
</tr>
<tr>
<td>5 kômã</td>
<td>kômkóm</td>
<td>'to dig'</td>
</tr>
<tr>
<td>6 kàñun</td>
<td>kàñwun</td>
<td>'to wither'</td>
</tr>
<tr>
<td>7 nàñ</td>
<td>mèñmrañ</td>
<td>'to break'</td>
</tr>
<tr>
<td>8 nì</td>
<td>nìfì</td>
<td>'to sit'</td>
</tr>
<tr>
<td>9 we</td>
<td>wìgwe</td>
<td>'to see'</td>
</tr>
<tr>
<td>10 rà</td>
<td>ðë</td>
<td>'to go in'</td>
</tr>
</tbody>
</table>

In table 6.21 the following devices are exemplified: (1) prefixation of ki-, (2) infixation of -p-, (3) prefixation and infixation, (4) infixation and ablaut (a→i), (5) reduplication of a monosyllabic verb, (6) reduplication of a disyllabic verb, (7) reduplication and ablaut, (8) reduplication and infixation, (9) reduplication, infixation and ablaut (a→i), (10) suppletion. Although most Kaingáng verbs are invariable, there are about 150 that have a plural form in one of the patterns illustrated in table 6.21. Even a verb that is probably a loan from Portuguese, paja 'to promenade' (Portuguese pasear [pasja]), pluralizes in one of these patterns, namely pìaja (infixation and ablaut).

Xavante (Central Jê) also expresses the number distinction of nominal arguments (S and O) by means of verbal agreement, but it systematically distinguishes three numbers, singular, dual and plural. About a score of verbs have three different stems, one for each number of S or O, such as sg. warà, ðu. alfabrò(i), pl. ðtjì'ìfìre 'to run'; sg. wì, ðu. ðjìmatìjiì, pl. ðjìhù 'to arrive'; sg. wìti(ri), ðu. parì, pl. ðjìbì 'to kill'; sg. bè(i), ðu. wàfìðì(ri), pl. tabrà 'to throw' (some stems have two allophones, one of them short, the other long, with the extra syllable put in parentheses in the examples above). In most verbs, however, the number distinction is made by means of a complex interplay of particles in the verb phrase.

5.4 Noun classification

The category of noun classes is not typical of Macro-Jê; these are only reported for the languages of the Kariri family. There are twelve prefixes that are attached to quantifiers and descriptive adjectives of dimension, consistency and colour; according to the shape of the noun's referents. The nouns themselves have no overt marking for class. According to Mamiani (1699), the twelve prefixes in Kipeá are the following: be- for hills, dishes, stools, foreheads, etc.; kro- for birds, stones, stars and round objects, such as beads, fruit, eyes, etc.; kru- for liquids and rivers; epru- for clusters and bunches; he- for sticks, legs and wooden objects; ho- for iron objects, bones and pointed things; mu-, mu- for edible roots; mu- for holes, wells, mouths, fields, valleys, fenced spaces; ro- for clothes, fabric and furs; waro- for roads, conversations, speeches, stories; bu- for houses, arrows, containers, corn-cobs and living beings (except birds), as well as for any other nouns not specified for the other prefixes.

The classifier prefix occurs with adjectives (which follow the head noun) and with quantifiers (which precede it), as exemplified in table 6.22.

There is a masculine/feminine gender distinction in 3rd person pronouns in the Southern Jê languages (i.e., Kaingáng and Xokleng). Rikbaktsa (Wiesemann 1986: 361) and Yaté (Lapenda 1968: 91) have a gender distinction in all three persons, and the latter language marks masculine and feminine genders on adjectives (see §6.6.), as well as on demonstratives (Meland 1968: 18, 19b).

5.5 Agreement marking on the verb

Many Macro-Jê languages have no agreement markers on the verb. Some of them, however, do mark the subject (most often, if it is third person) in agreement with the corresponding nominal phrase in the clause. Kaingáng (Jê) exemplifies the complete absence of markers in (12) and (13), whereas Kipeá (Kariri) shows 3rd person and 1st inclusive agreement in (14) and (15):

(12) ?ip repre wì jê  
1sg brother S stand
'My brother is standing.'

Table 6.22 Class agreement in quantifying and qualifying phrases in Kipeá

<table>
<thead>
<tr>
<th>spherical objects</th>
<th>conical objects</th>
<th>sinuous objects</th>
<th>convex objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ufe 'sun, day'</td>
<td>udza 'knife'</td>
<td>wo 'snake'</td>
<td>pika 'stool'</td>
</tr>
<tr>
<td>'long'</td>
<td>ufe kro-ji</td>
<td>udza ja-ji</td>
<td>wo ho-ji</td>
</tr>
<tr>
<td>'bright'</td>
<td>ufe kro-dzodo</td>
<td>udza ja-dzodo</td>
<td>wo ho-dzodo</td>
</tr>
<tr>
<td>'one'</td>
<td>kro-bihe uje</td>
<td>ja-bihe udza</td>
<td>ho-bihe wo</td>
</tr>
<tr>
<td>'many'</td>
<td>kro-jo uje</td>
<td>ja-jo udza</td>
<td>ho-jo wo</td>
</tr>
</tbody>
</table>
(13) mĩŋ wi kāgįren tāŋ tī
The jaguar killed a tayra (T. barbara).

(14) more si-te karai
'The white man comes soon.'

(15) ku-te di
'We (incl.) will come.'

Sorne languages use a marker on the verb only if the subject nominal phrase does not immediately precede an intransitive verb, for example Maxakalí in (16) and (17) (Pereira 1992: 83):

(16) pitjapep tjihep
'The duck arrives.'

(17) tį-tjihep pitjap
duck arrive 3-arrive duck
'The duck arrives.'

The same marking occurs in Maxakalí for the direct object of transitive verbs, as in (18) and (19) (Pereira 1992: 88), and is fundamentally the same device as that shown in §5.1 above as the marking of non-contiguity versus contiguity of a determiner.

(18) tik te tį-tįsit kipiįtk
man ERG 3-sharpen axe
'The man sharpens the axe.'

(19) tik te kipik tįst
man ERG axe sharpen
'The man sharpens the axe.'

An example with transitive verbs in Timbira (Popjes and Popjes 1986: 163) is (20):

(20) jàko te po pupun ne iįt-kuran
Jacó ERG.PAST deer see and 3-kill
Jacó saw the deer and killed it.

Yaté, Karajá and Guató have more complex verb morphologies and manifest verb agreement not only with the subject, but also with the object. Guató is particularly complex; it has not only personal prefixes like other Macro-Jé languages, but also personal suffixes. Moreover, Guató verb morphology is a compound of 'three different inflectional patterns: one for first person singular, a neutralization of systems; another for second and third persons singular, which follows an ergative/absolutive system; and still a third for plural, which behaves according to a nominative/accusative system' (Palácio 1986: 369). Palácio’s chart 6 is reproduced here as table 6.23. (Note that non-singular covers 1 dual exclusive, 1 plural inclusive and 1 plural exclusive.)

Examples of this person marking system in a transitive verb are the following:

(21) na-gą-bagákti-tjįjıo
INDIC-2A-hit-lsgO
You hit me.

(22) n(a)-e-bagáki-he
INDIC-3sgA-hit-20
'He hits you.'

It is probable that the object prefixes ge- and dʒe- are the result of a contraction of the expected sequences e-geo- and e-dʒe- (cf. Palácio 1984: 68; 1986: 370). Third person prefixes are obligatory even when the corresponding nominal phrases are fully expressed in the clause, as in (23); therefore they are truly agreement markers.

(23) n(a)-e-bagákti-ə
INDIC-3sgA-hit-3sgO
DET-man DET-woman
The man hits the woman.'

6 SYNTAX

6.1 Constituent order in declarative sentences

The languages of the Jé family have SV and AOV as the most frequent constituent orders in their declarative clauses. Examples (24) and (25) from the Canela dialect of Timbira illustrate these orders for one-argument and two-argument verbs respectively.
The same pattern also prevails in Maxakalí, Krenák, Yaté, Karajá, Boróro and Rikbaktsá.

Data on Kamaka, Purí and Ofayé are so limited that no statement on their clause syntax is possible. This is an irremediable situation for the languages of the first two families, which are now dead, but it is not so for Ofayé, from whose last survivors (about a score) linguistic data could still be recorded.

Kariri and Guató represent strong departures from the pattern illustrated above. In both language families clauses are normally verb-initial. In Kipeá of the Karirí family and in Guató, clauses with one-argument verbs show the same order VS, but those with two-argument verbs differ in the relative position of the arguments. In Kipeá (see examples 26 and 27) we get VOA (where the A is marked by the ergative preposition no). In Guató we find VAO, as shown in (28) and (29):

(26) si-te karai
3-come white.man
'The white man comes.'

(27) si-pa kradzo no karai
3-kill cow erg white.man
'The white man kills the cow.'

(28) na-kini g-eti
INDIC-sleep DET-boy
'The boy sleeps.'

(29) ma-e-ro g-tpagu g-tki
IMPERFv-3-eat DET-jaguar DET-rabbit
'The jaguar ate the rabbit.'

6.2 Adpositional phrases

All Macro-Jê languages except Karirí and Guató have postpositions and not prepositions, as may be seen in the following examples.

Kayapó
(30) pur kâm ba a mês pijo re
garden in lsg you to fruit pick
'I pick up fruits for you in the garden.'

Maxakalí
(31) piticnãg Ti-cip mim ti
bird 3-be tree on
'The bird is on the tree.'

The Karirí languages and Guató, which have verb-initial clauses, have prepositions and not postpositions:

Kipeá
(32) pa kri do udza
kill PERFv with knife
in DET-river
'It was killed with a knife.'

Guató
(33) jo go-dékê
in DET-river
As seen in the last four examples, some adpositions have different allomorphs when they follow a pronoun or pronominal prefix.
6.3 Genitive phrases

With the exception of Karirí and Guató, the Macro-Jê languages have the possessor preceding the possessed noun (with no overt marker of possession):

\[
\text{Timbíra} \quad \text{Maxakalí}
\]

(39) kapi k'ra ?i tit fikpot
Capi child 3 mother grave
'Capi's child' 'his mother's grave'

Guató and Karirí are again the languages that behave differently, having an inverted order for their genitive phrases. In these two Macro-Jê families the heads of noun phrases precede their determiners (or dependents). If the relation between head and determiner is one of possession, the head is inflected for 3rd person (or non-contiguous) possession as shown in the following examples from Guató:

(41) i-pána g-ák'o
3-tail DET-monkey
'the tail of the monkey'
(lit. 'its tail the monkey')

Note that Kipeá has the reverse order in compounds:

(43) s-era ware (44) boro-po
3-house priest arm-eye
'the house of the priest' 'elbow'

Many languages of the Macro-Jê stock distinguish alienable from inalienable possession. In general, alienable possession is expressed by means of an inalienable generic noun, meaning 'thing' or 'belongings' or 'possession'. In some languages there are two or more generic nouns, distinguishing classes of possessable things. The Jê languages Timbíra, Kayapó and Panará have only one generic form, -ô, as in examples (45) and (46) from the Canela dialect of Timbíra.

(45) kapi j-ô pur
Capi cnt-POSSESSION field
'Capi's field'

(46) h-ô wapo
NCNT-POSSESSION knife
'his knife'

Boróro has two generic nouns for expressing alienable possession, -aku 'domesticated animal' and -ô 'thing', as illustrated by examples (47) and (48).

(47) e-aku kogariga
3-DOmESTIC-ANIMAL chicken
'his chicken' (lit. 'his-domesticated-animal chicken')

(48) e-o tori
3-thing stone
'his stone' (lit. 'his-thing stone')

In Kipeá, of the Karirí family (see §5.4), there are twelve generic nouns which refer mainly to as many different classes of alienably possessable objects, including animals, according to the way in which they were acquired. The generic noun is possessed in the same way as any other noun and is linked to the specific noun of the possessed object by means of the preposition do. Three generic nouns are illustrated here:

(49) hi egki do kradzo
1sg DOMESTIC-ANIMAL of cow
'my (raised) cow'

(50) dz-uapru do murawo
1-game of wild.pig
'my (hunted) wild pig'

(51) dz-uito do udza
1-find of knife
'my (found) knife'

According to the way in which its referent was acquired, a noun may occur with different generic nouns, as in examples (52)–(54).

(52) hi egki do sabuka
1sg DOMESTIC-ANIMAL of chicken
'my (raised) chicken'

(53) dz-ukisi do sabuka
1-portion of chicken
'my chicken (received in a partition of goods)'
With the exception of *egki* 'domesticated animal', all possessable generic names of Kipeá begin with *u*; this *u* is probably a morpheme cognate with Boróro *o* 'domesticated animal', Timbira *-o* 'belongings', Maxakali *jo - jo* 'belongings', and probably also Karajá *ad* 'domesticated animal'.

6.4 Demonstrative phrases

Demonstratives follow the head noun in some languages and precede it in others. Canela (of the Jé family) and Krenák illustrate the first situation, as in example (55), whereas Kipeá of the Kariri family, Boróro and Guató show the reverse, as in example (56).

\[
\begin{align*}
\text{Canela} & \quad \text{Kipeá} \\
(55) & \quad \text{ig i era} \\
& \quad \text{this dog} \\
& \quad \text{this house} \\
& \quad \text{'this dog'} \\
& \quad \text{'this house'}
\end{align*}
\]

6.5 Numeral phrases

In the Kariri and Guató families numerals precede the head noun, whereas in the other language families they normally follow it, as in examples (57)–(61). In the Kariri languages the numerals carry a classifying prefix in agreement with the class of the phrase head (see §5.4 above), as in (57), but they take no prefix if the head noun is unclassified, as in (58). In Yalet the numeral for 'one' forms a compound with the noun, and this compound is marked for gender as in (60).

\[
\begin{align*}
\text{Kipeá} & \quad \text{Guató} & \quad \text{Yaté} \\
(57) & \quad \text{bu-bihe erumu} & \quad \text{bihe tupa} & \quad \text{tfáji kaka-ne} \\
& \quad \text{CL-one squash} & \quad \text{one god} & \quad \text{man good-MASC} \\
& \quad \text{‘one squash’} & \quad \text{‘one god’} & \quad \text{‘a good man’} \\
(62) & \quad \text{fika kaka-o} & \quad \text{udza ja-fí} & \quad \text{g-óda g-ítavi} \\
& \quad \text{man good-MASC} & \quad \text{knife CL-long} & \quad \text{DET-basket DET-heavy} \\
& \quad \text{‘a good man’} & \quad \text{‘a long knife’} & \quad \text{‘the heavy basket’}
\end{align*}
\]

6.6 Adjectival phrases

In Macro-Jé languages the adjective follows the noun. This is true of the languages with constituent order AOV as well as of those with orders VOA and VAO. Only in the Yalet family is there gender agreement of the adjective with the noun, see (62) and (63). In the Kariri family the small set of adjectives of dimension, colour, and quantity shows class agreement with the head nouns, as in (64)–(66) from Kipeá (see §5.4 above). In Guató the adjective takes the same marker for determination as its head noun, as in (67).

\[
\begin{align*}
(62) & \quad \text{fika kaka-o} & \quad \text{ufa} & \quad \text{udza ja-ne} \\
& \quad \text{man good-MASC} & \quad \text{CL-many sun} & \quad \text{knife CL-sharp} \\
& \quad \text{‘many days’} & \quad \text{‘a long knife’} & \quad \text{‘a sharp knife’} \\
(65) & \quad \text{ufa} & \quad \text{udza ja-ne} & \quad \text{g-óda g-ítavi} \\
& \quad \text{CL-many sun} & \quad \text{knife CL-sharp} & \quad \text{DET-basket DET-heavy} \\
& \quad \text{‘many days’} & \quad \text{‘a long knife’} & \quad \text{‘the heavy basket’}
\end{align*}
\]

6.7 Ergativity

There are some techniques for linking S and O that are universal and do not indicate that a language in which they occur has ergative properties. For instance, if a verb marks the number of a core argument this will always be S in an intransitive and O in a transitive clause, as reported for Kaingang and Xavante in §5.3.

There is ergativity in Jé languages where an A NP can, at least in some contexts, take a preposition that appears to have ergative function, as in the following example from Timbira:

\[
\begin{align*}
(68) & \quad \text{ rop kakbwin} \\
& \quad \text{1sg ERG + PAST dog beat} \\
& \quad \text{‘1 (recently) beat the dog.’}
\end{align*}
\]
Aryon D. Rodrigues

In Maxakalí any transitive sentence has its subject marked the ergative post-
position te, as in examples (69) and (70).

(69) tik te kipiktfit
man ERG axe sharpen
'The man sharpens the axe.'

(70) kipik te mǐm kaʔok mep
axe ERG wood hard cut
'The axe cuts hard wood.'

Kipeá (Karírí family) has VS, VOA constituent order with the A argument
marked by an ergative preposition no:

(71) peho i-wo dzu mo imera
flow 3-way water in field
'The river flowed on the fields.'

(72) sō hietsā no wo
bite 1sg ERG snake
'A snake bit me.'

This now-extinct language was described by Father Mamiani (1699) who struggled
to fit it into a classical grammatical framework. He characterized all transitive verbs
as inherently 'passive' (with no corresponding active form).

But, besides the semantics of the verbs, there is a morphosyntactic device that
reveals the ergative character of Mamiani's 'passive' verbs. Every Kipeá verb may
derive a nominalization of its absolutive (S or O) argument (the 'subject' according
to Mamiani), by means of the absolutive prefix di- and the nominalizing suffix
-ri, as in the following examples:

(73) di-te-ri
ABS-come-NMLZR
'the one who comes'

(74) di-pa-ri
ABS-kill-NMLZR
'the one that was killed'

(75) udza di-di-ri no ware
knife ABS-give-NMLZR by priest
'the knife that was given by the priest'

Only Mamiani's 'passive' verbs, however, have a second nominalization, referring
to the agent of the action, therefore the A argument. This nominalization takes the
ergative prefix du- and the same nominalizing suffix -ri, as in the following examples:

(76) ware du-di-ri udza
priest ERG-give-NMLZR knife
'the priest who gave the knife'

(77) koho du-nio-ri arākie
that ERG-make-NMLZR sky
'he was who made the sky'

There is further discussion of the ergative character of Kipeá in Larsen (1984),
who also indicates an SO/SO syntactic pivot for subordinate clauses. (Larsen's discus-
sion also suggests a class of 'extended intransitive' verbs in Kipeá.)

6.8 Valency-changing processes

6.8.1 Reflexives and reciprocals

Both morphological and syntactic devices are found in the Macro-Jê languages for
reflexives and, in some cases, also for reciprocals (not every language clearly distin-
guishes these two processes). Morphological devices may be seen as the derivation
of an intransitive verb from a transitive one, whereas syntactic devices involve a
reflexive or reciprocal use of a generic pronoun (and maintain transitivity). This is
the case with languages of the Jê family, such as Timbira (Canela dialect) with the
generic reflexive object amji and the reciprocal ajpēn:

(78) i te amji pitar
1sg PAST REFL defend
'I defended myself.'

(79) ku-te amji mā pī jakhēp
3-ERG.PAST REFL for wood cut
'He cut wood for himself.'

(80) jakē me kapi aipēn mā pī jakhēp
Jacó and Capi RECIP for wood cut
'Jacó and Capi cut wood for each other.'

Xavante, a Central Jê language, has a reflexive prefix tsįtį- ṭmaDa 'to look at', tsįtį-
ṭmaDa 'to look at oneself, to take care of oneself'. Analogously, Yate has a prefix
sa-tu/tu 'to cut', sa-tu/tu 'to cut oneself'.

6.8.2 Causativization

Some languages in the Macro-Jê stock have morphological devices for the forma-
tion of causative verbs, whereas other languages have only syntactic means of
causativization, involving verbs such as 'to make' or 'to cause'. Morphological devices consist essentially in deriving a transitive stem from an intransitive one. Among the languages that have morphological derivation of causatives, some can causativize only intransitive verbs while in others they can also derive transitive verbal stems from nominals.

The Jé language Timbirá causativizes syntactically by means of the verb *na* 'to make', to which the causativized verb is linked by the particle *na* (cf. Popjes and Popjes 1986: 142-3):

(81) kapi te i jôt na i to
Capi ERG.PAST 1SG sleep na 1SG make
'Capi made me sleep.'

(82) i te i prô j-apen na o-ton
1SG ERG.PAST 1SG wife CNT-work na CNT-make
'I made my wife work.'

An interesting example of a syntactic causative involves Boróro. In this language there are two causative morphemes, *d3* 'to cause' (simple causative) and *g3* 'to cause to begin' (inceptive causative):

(83) a re boe e ridiwa-d3
you NMLZR people they KNOW-CAUS
'You caused people to know.'

(84) a-re boe e ridiwa-g3
you NMLZR people they KNOW-INCAUS
'You caused people to begin to know.'

*d3* and *g3* are not suffixes as examples (83) and (84) might suggest. They are clitics that occur after intransitive verbs (as above), but also after the subject of causativized transitive verbs, as in (85):

(85) i re a-d3 bola barigu imedi ae
1SG NMLZR you-CAUS ball throw man to
'I caused you to throw the ball to the man.'

Looking now at morphological devices, some languages use prefixation and others exhibit suffixation. Kipeá (Karirí) has a causative prefix *mi-* added to intransitive verbs: *pere* 'to go out', *mi-pere* 'to cause to go out'; *te* 'to come', *mi-te* 'to bring'. Kaingâng (Jé) has a suffix *-n* which occurs with both intransitive and transitive verbs: *fi* 'to be old', *fi-n* 'to cause to be old'; *râ* 'to be ripe', *râ-n* 'to make ripe'; *we* 'to see', *we-n* 'to show'.

6.9 Switch-reference

Some Macro-Jé languages distinguish coordinate clauses with the same subject from those with different subjects. In Kipeá this distinction is made by using different prefixes on the verb of the coordinated clause. When the subject of this clause has the same reference as the subject of the first one, it is marked by *di-* whereas the non-contiguity marker *si-* is used for subjects differing from that of the first clause (Mamiani 1698: 61, 49):

(86) doro si-te bo arâkke no radâ do di-wi do then 3-come from heaven in earth for SS-become to
_tsoôho do di-nja nodehê_ people for SS-die also

'Then he came from the heavens to the earth in order to become people and also to die.'

(87) mo s-unu-te Adam si-pei-kri i-mesu
in CNT-sleep-MLZR Adam NMLZR-take.out-PERFV NMLZR-rib
no tupâ i-bo bo si-nio i-bujêwoho Eva
ERG God NMLZR-from for NMLZR-make NMLZR-body Eve

'During Adam's sleep God took out his rib for making Eve's body.'

Another switch-reference strategy is found in some Jé languages and in Maxakali. Different connective particles or conjunctions distinguish clauses with different subjects from those with the same subject, as shown in examples (88) and (89) for the Canela dialect of Timbirá (Popjes and Popjes 1986: 147) as well as in (90) and (91) for Maxakali (Popovich 1986: 355).

(88) kapi te po kurun ne ke ha ku-kua
Capi ERG.PAST deer kill and 3.SS FUT 3-eat
'Capí killed a deer and will eat it.'

(89) kapi apu ajkahu mâ hitsi apu nô
Capi CONT run and.DS his.wife CONT lie.down
ne nôr and.SS sleep
'Capí is running and his wife is lying down and sleeping.'

(90) 7i-mông ti ?-nîn
3-go and.SS 3-come
'He went and returned.'
The Macro-Jê stock of languages has so far been a mere hypothesis, or rather a phonological equations involved. The spatial distribution of its members is very large and the lexical bundle of hypotheses, about a possible remote common origin for the languages and a large area of linguistic spread or whether Karirí and Guató may be viewed as independent examples of the same sort of typological rearrangement. To judge from the grammatical survey in §5 and §6 above, even though rather superficial, shows that there are striking typological similarities among all the twelve branches of the Macro-Jê stock. It shows also that two branches, Karirí and Guató, nearly diverge in syntactic typology from the other ten branches; whereas the latter are predominantly verb-final and postpositional, with genitive-noun phrases, Karirí and Guató are both verb-initial and prepositional, with no genitive phrases. It is remarkable that Karirí and Guató are not geographical neighbours but, on the contrary, Karirí lies in the northeast of the Macro-Jê area, on the lower São Francisco river, whereas Guató is found in the extreme southwest of the area, on the Paraguay river. As yet it is too early to decide whether these constitute cases of marginal conservatism in internal evidence in Karirí, the rearrangement hypothesis is more likely for this family, since its languages show residual traces of the typological alignment prevailing in other Macro-Jê languages (e.g., postpositional use of the prepositions when their objects are pronominalized, genitive-noun order in compound words, see §6.2 and §6.3 above).

7.2 Phonological equations

The Macro-Jê stock of languages has so far been a mere hypothesis, or rather a bundle of hypotheses, about a possible remote common origin for the languages involved. The spatial distribution of its members is very large and the lexical properties can be undertaken. The above survey of selected aspects of some of the better-known languages may give an idea of the kind of structural phenomena manifested in this huge language group and will hopefully stimulate research in the languages that are still alive, most of which have been and still are—exposed to strong adverse sociocultural pressures and are seriously endangered.

7 ABOUT THE CONSISTENCY OF MACRO-JÊ AS A GENETIC GROUP

7.1 A brief appraisal of the grammatical affinity

The grammatical survey in §5 and §6 above, even though rather superficial, shows that there are striking typological similarities among all the twelve branches of the Macro-Jê stock. A sample of these correspondences and equations is presented here in order to enable the reader to have an idea of the likelihood of the genetic relationship of the languages involved. Table 6.24 displays the phonological correspondences and the numbers for the examples presented following the table.

Table 6.24 Phonological correspondences among the Macro-Jê language families

<table>
<thead>
<tr>
<th>Language families</th>
<th>proto-MJ lexical items</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>p</td>
<td>w</td>
</tr>
<tr>
<td>m/p</td>
<td>w</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>k/θ</td>
<td>k</td>
</tr>
<tr>
<td>kr</td>
<td>her</td>
</tr>
<tr>
<td>gr</td>
<td>gr</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

1. arm: la pā, lb pa, lc pa, ld pē, IV po, VI bo, VII fe 'armpit', IX pc, XI pē
2. arrow: lc po, ld pun, II wāj, III poj, V pan, pan, VI buj-ku, VIII wāji, X bēja "bow"< *bēja-ika "arrow-bow"
3. axe: lb kā-mep, III kā-pik, IV kā-pok, V kā-mag
4. back: VI wari, VIII bāro, IX -hor, X pori
5. caiam: lb mā-tii, II wej-e, III maʔi, mā, X wai
6. come: lb rē, lc ri 'to go', II ni, III nin, IV ne, V ni, VI te, VII th
7. drink: lb kōm, bō, III tfo-top, tjan, IV tfo, V some, VII kō, VIII ō, X ku, XI dē, XII ku
8. dry: lb gara, lc ʔr, VI kra, X kiwē
9. eat: Id raγ 'to swallow', VI do, VIII ra 'to eat meat', IX rō, XI ro
I 26 night: §i" ~; (I
10 egg: lb gø, lc ?re, ld kr̆, ll sa-k-ke, III kic, VIII ði, IX kíte, XI kíø, XII kare
11 eye: lb na, ta, lc ta, II to, VII øø, VIII ace
12 foot: Ia pena, lb par, lc para, Id pøn, II wade, III patu, IV po, VI hé, biri-, VII jé-he, fet-, VIII wa, IX par, X bire, XI òbò, XII piri
13 give: lb gø, lc tsø, III høm, IV -úp, hum, VII ko, VIII ø, IX no
14 hair: Ia fe, lb ki, kø², II ke, III tfe, IV ke, V ke, tfe, XI ki
15 head: Ia kr̆, lb kr̆, lc ?ra, ld kr̆, II hero, xaro, IV kren, VIII ra, IX kíte, XII -hara-
16 hear: lb ma, lc wa-pa, Id mëg, IV paw, IX pæj, X mea-ríd
17 honey: lb meq, Id møg, III paq, IV paq, IX pík, XI pagua
18 I: lb i, lc ò, Id òn, III ?k, VI hé, VII i, X i, XII ik
19 in: ld kí, VII ke, VIII ki, X gi
20 leaf: Ia arø-tife, II ere, VI æræ, X arí 'leaf', aro 'small leaves'
21 liver: lb ma, lc pa, Id tø-mè, III ta-ma-qaï, VI ba, IX pa, XI pe
22 long: lb ri, Id ríu 'to push', II roro, IV røn, VIII rehe, IX ra, X rai-
23 re, XII (ze-ze)
24 maize: V maki, maeki, VII masã/kamasti, VIII maki, XII na-fi
25 mount: ld kr̆, II heri, kere, III pø-køi
26 night: lc børa, II børa, V meri, IX wez
27 penis/male: Ia c̆, VI ca 'male'
28 possession: lb ø, III ø, òp, jøg, VI w-, X ò
29 short: Id rør, VII lujiñ, X ro-gu
30 sing/dance: Ib gø, lc ò-re, 'dance', Id gøn 'dance', II gøre, III køj, IV gøø, V gøre, VII køø-øø, VIII ðe, IX kir̆h, XII kari 'to dance'
31 skin/bark: lb ka, lc ha, II ka, III sjøj, kaj, IV kat, VII køø-øø, IX ha, X -ka
32 sleep: Ia ejø, lb gør (nø, nør 'to lie'), lc jødø (dø 'to lie'), Id nør, II jøndø, høndø, III pøn, høm, ?øm, VI uu, VIII ðø, IX no, nør 'to sit', X mødø, XII uu
33 stone: Ib køn, kën, lc ?ødø, II kéø, VI kro, IX køtø
34 stones: Ia pø, VII fova, XI fø 'ground'
35 sun: lb nì, lc bøla, VI bøt 'star', VII føøa, X meri
36 to: lb nì, lc bø, Id nì, VII ma, VIII bø
37 walk/go: lb mømør, lc bøbøi, Id mø 'go pl.', II maq, mán, III møg 'to go', IV møg, mø, V mu, mon 'to go', VI wo, X meru

The proto-phonemes (marked with an asterisk) to the right of each series of phonological correspondences in table 6.24 are very tentative. There remain many inconsistencies in the correspondence of vowels and consonants across the twelve Macro-Jé families, and the relative scarcity of data makes it very difficult to find examples that would substantiate each series of possible cognate words. Another difficulty, well known in historical linguistics, comes from the fact that most comparable words or morphemes in these languages are very short ones, consisting in general of only one or two syllables and therefore without the phonological redundancy that makes a genetic connection really plausible. However, there is a reasonable consistency in a good part of the series of correspondences that go across the whole Macro-Jé stock. Moreover, most of the meanings in these series belong to those sections of the lexicon (universal concepts) that are less exposed to cultural influences, and, therefore, are less likely to be loans from one language to another. These factors make us reasonably confident that the Macro-Jé hypothesis – namely the hypothesis of a common genetic origin for the Macro-Jé languages – may be proved correct in the future – if not for all the language families included here, at least for most of them.

BIBLIOGRAPHY

1976. 'Nasality and the internal structure of segments', Language, 52:326-44.
A listing of the various languages in the Tucano language family that are currently in use first requires a decision as to which languages are sufficiently distinct to warrant a separate listing. Thus it may be noted that some entries in table 7.1 have two names. Barasano (also known as Panera) and Taiwano (also known as Eduria) differ mainly in pitch-stress on words (Jones and Jones 1991: 2) and so are grouped together. Retuara and Tanimuca differ mainly in a few lexical items (Strom 1992: 1). Although I consider Bará to be distinct from Waimajá, there is not sufficient data on Bará to list it separately. Pisamira would be included if there were sufficient data available. Waltz and Wheeler (1972: 128) group Pisamira with Tuyuca and Bará under the name Pápiwa, and indeed a taped word list that a Tuyuca lady and I listened to indicates that it has much in common with Tuyuca.

The Eastern and Central Tucano languages are spoken in the northwestern Amazon Basin in the state of Vaupés, Colombia, and in the state of Amazonas, Brazil. The Western Tucano languages are spoken in southwestern Colombia along the Putumayo and Caquetá rivers, and along the Putumayo and Napo rivers in Ecuador and Peru. In table 7.1, the letters E, W and C represent Eastern, Western and Central. The subgroupings in table 7.1 are taken from Waltz and Wheeler (1972), and are based on phonological and lexical affinities between the Tucano languages. Malone (1987), on the basis of her study of the development of the current languages from the proto-language, has the same groupings, with the exception that she pulls E6 and E7 out of the Eastern category and puts them in a fourth major category, tentatively also including E11 and E12 in that category. The approximate locations of the Tucano languages are shown on map 6.
The Tucano groups are rapidly becoming acculturated to many aspects of the dominant Latin culture, yet many of the people in these groups still carry out their daily life much as their ancestors did: using slash-and-burn agriculture, fishing, and eating casabe bread with a fish sauce prepared with hot peppers. Individual family houses with corrugated zinc roofs are replacing the 'maloca', a large communal house typically shared by a man and his sons and their families.

In general, the Eastern Tucano language groups are exogamous with respect to language, and individuals continue to speak their own (i.e., their father's) language, as long as those they are speaking to understand it. Thus, children grow up knowing both their mother's and their father's language, as well as having at least a passive competence in the languages of the other women in the village or longhouse. For more information on this unique multilingual situation, see chapter 14 of this volume as well as Sorensen 1967.

Word lists taken in the 1960s and before indicate a very high degree of shared vocabulary between the languages. More recent studies of verb morphology and discourse grammar indicate that the actual degree of intelligibility may not be as high as the lists of shared vocabulary might indicate (see West 1977).

---

Table 7.1  
The Tucano language family (B, Brazil; C, Colombia; E, Ecuador; P, Peru)

<table>
<thead>
<tr>
<th>Region</th>
<th>Language</th>
<th>Population</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>WESTERN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western north</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 Koreguaje</td>
<td>2,000</td>
<td>Ríos Orteguaza and Caquetá (C)</td>
<td></td>
</tr>
<tr>
<td>W2 Secoya</td>
<td>400</td>
<td>Río Putumayo (E)</td>
<td></td>
</tr>
<tr>
<td>W3 Siona</td>
<td>300</td>
<td>Río Putumayo (C, E)</td>
<td></td>
</tr>
<tr>
<td>Western south</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W4 Orejón</td>
<td>200-300</td>
<td>Ríos Napo, Ampí Yacú and Algodón (P)</td>
<td></td>
</tr>
<tr>
<td>CENTRAL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1 Cubeo</td>
<td>6,000</td>
<td>Ríos Vaupés, Cuduyari, Querari and Pirabotón (B, C)</td>
<td></td>
</tr>
<tr>
<td>C2 Tanimucal/Returá</td>
<td>300</td>
<td>Ríos Guacayá, Mirití, Oiyaká and Apaporis (C)</td>
<td></td>
</tr>
<tr>
<td>EASTERN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern north</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1 Pirapataya</td>
<td>1,100</td>
<td>Río Papuri (B, C)</td>
<td></td>
</tr>
<tr>
<td>E2 Tucano</td>
<td>4,100-4,600</td>
<td>Río Papuri and Caño Paca (B, C)</td>
<td></td>
</tr>
<tr>
<td>E3 Wanano</td>
<td>1,100</td>
<td>Río Vaupés (B, C)</td>
<td></td>
</tr>
<tr>
<td>Eastern central</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4 Bará/Waimajá</td>
<td>500-600</td>
<td>Caños Colorado, Fríjol, Lobo, Inumbú and Yapú, Río Tiquié (B, C)</td>
<td></td>
</tr>
<tr>
<td>E5 Carapana</td>
<td>600</td>
<td>Caño Ti, Ríos Piraparaná, Papuri and Vaupés (C)</td>
<td></td>
</tr>
<tr>
<td>E6 Desano</td>
<td>1,000</td>
<td>Ríos Papuri and Vaupés (C)</td>
<td></td>
</tr>
<tr>
<td>E7 Sirianó</td>
<td>250-300</td>
<td>Ríos Paca and Viña (C)</td>
<td></td>
</tr>
<tr>
<td>E8 Tutuyo</td>
<td>350</td>
<td>Ríos Piraparaná and Papuri, Caño Yapú (C)</td>
<td></td>
</tr>
<tr>
<td>E9 Tuyuca</td>
<td>725</td>
<td>Ríos Papuri and Tiquié, Caño Inumbú (B, C)</td>
<td></td>
</tr>
<tr>
<td>E10 Yurutí</td>
<td>200-250</td>
<td>Caños Paca and Ti (C)</td>
<td></td>
</tr>
<tr>
<td>Eastern south</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11 Batañano/Taiwano</td>
<td>250</td>
<td>Río Piraparaná (C)</td>
<td></td>
</tr>
<tr>
<td>E12 Macuna</td>
<td>350</td>
<td>Ríos Cometa, Piraparaná and Apaporis (C)</td>
<td></td>
</tr>
</tbody>
</table>

---

2  PHONOLOGY

The proto-consonants according to Malone (1987: 7) are given in table 7.2. She posits nine consonant phonemes plus suprasegmentals of nasality, tone and/or stress, and glottal stop.

The sixproto-vowel system is given in table 7.3.
Table 7.2  Proto-consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>alveolar</th>
<th>velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>*p</td>
<td>*t</td>
<td>*k</td>
</tr>
<tr>
<td>voiced stop</td>
<td>*b</td>
<td>*d</td>
<td>*g</td>
</tr>
<tr>
<td>voiceless sibilant</td>
<td>*s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced semi-vowel</td>
<td>*w</td>
<td></td>
<td>*j</td>
</tr>
</tbody>
</table>

Table 7.3  Proto-vowels

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>*i</td>
<td>*i</td>
<td>*u</td>
</tr>
<tr>
<td>low</td>
<td>*e</td>
<td>*a</td>
<td>*o</td>
</tr>
</tbody>
</table>

The present-day languages all maintain the voiced/voiceless contrast in the stops. Koreguaje has nasals rather than voiced stops plus both unaspirated and aspirated voiceless stops. Wanano has three sets of stops: voiced, voiceless unaspirated and voiceless aspirated. Orejón, in addition to voiced bilabial and alveolar stops, has voiced implosive stops at those points of articulation. Secoya and Siona have a combination of nasals and voiced stops in contrast to the voiceless stops.

2.1 Syllable structure

The basic syllable structure for the Tucano languages is (C)V. For those languages with a glottal stop, there is an additional syllable type: (C)V?. W1–3 and C2 have diphthongs, resulting in the additional syllable type (C)VV for those languages. The other languages have geminate vowels which are generally analysed as a sequence of two syllables – CVV – due to accent placement.

Examples from Tuyuca and Tucano include:

(1) dií 'blood'  dií 'clay, meat'  (Tuyuca)
    dií 'blood'  diʔi 'clay, meat'  (Tucano)

2.2 Segmental phonology

All of the Tucano languages, with the exception of Tanimuca, have a six-vowel system which is the same as the proto-vowel system given in table 7.3. Tanimuca has a five-vowel system, lacking the high central unrounded vowel /i/.

The glottal stop occurs not only in E1 and E6, but also in E2, E3, C2 and all the Western languages. In W1 the glottal stop is very strong. In W2 and W3 it is very weak. There is no information available on W4. The glottal stop of C2 is strong, as it is in E1, E3 and E6. In E2 its effect is to lengthen and laryngealize the vowel that precedes it.

Siona (W3) and Koreguaje (W1) each have eighteen consonant phonemes. Wanano (E3) has sixteen. The rest of the languages have between ten and twelve consonant phonemes as follows:

E1, E2, E6, W2 and W4 have twelve;
C1, C2, E5, E7, E9, E10 and E12 have eleven; and
E4, E8 and E11 have ten consonant phonemes.

In the Eastern and Central languages /b, d, g, r, w, j, h/ are realized as /m, n, ŋ, ŋ, ġ, w, j, h/ in nasal morphemes. The flap /r/ is realized as a lateral flap in some environments. The glide /j/ varies from a nonsyllabic vowel to an alveo-palatal affricate [dʒ] in oral morphemes, and from a nasal glide to a nasal in nasal morphemes.

2.3 Nasalization

In the Central and Eastern Tucano languages, nasalization is an autosegment associated with the morpheme (see Barnes 1996). Morphemes are marked in the lexicon as nasal, oral, or are unmarked. Nasalization spreads progressively in all the languages to unmarked morphemes. In Barasano, Desano and Siriano nasalization also spreads regressively, affecting only a few specific morphemes (see Miller forthcoming). A typical example of nasalization spreading through an unmarked morpheme is given in the following example from Tuyuca, using the unmarked imperative suffix /-ja/.

(2) bia-ja [bija] close-IMPER 'close (it)'

Malone, through her study of proto-Tucano (p.c.), has concluded that Tanimuca resulted from a group of Arawakan speakers who learned a Tucano language. This would help explain why Tanimuca has only the five vowels found in Arawakan languages. In addition, she has said that it can be useful to include Tanimuca when comparing Tucano morphemes, but it is not very useful to include Tanimuca when doing a comparative study of the grammars of Tucano languages. Note the number of instances in the grammar section where Tanimuca differs from all the other languages therein described.
The inventory of consonants through which nasalization spreads varies from language to language, but most of the languages include the liquids /w, j, h/ and the flap /l/ in their inventory, if they have suffixes that begin with those consonants.

2.4 Accent

The Tucano languages have accent, or pitch-accent systems in which there is high pitch vs low pitch. High pitch is shown by ('). and low pitch is unmarked. High pitch is generally associated with accent. Some of the languages, such as Desano, Tuyuca, Yurutí and Cubeo, have one accented syllable per phonological word. Several of the languages, among them Barasano, Carapana, Macuna, Siriano, Tucano and Waimajá, may have two or more contiguous high-pitched syllables, all of these syllables being equally accented. In Secoya, words of four or more syllables have accents on alternating syllables. Examples of these are:

- one accented syllable per phonological word in Cubeo (Salser, Morse and Ferguson, 1986: 3):
  (4) hatióó 'she is cooking'
- multiple high pitch in Barasano (Jones and Jones 1991: 11-12):
  (5) ságáó 'thigh'
  wasóó 'robe'
- accent on alternating syllables in Secoya (Johnson and Levinsohn 1990: 19):
  (6) de?wáó'tá?adé 'in the place where he fixed it'

3 MORPHOLOGY

The Tucano languages are almost entirely suffixing languages, with a strong tendency towards agglutination. One exception is the prefix ka-, the specific nominal referent, possibly borrowed from Arawak, that occurs on nouns in E4, E5 and E8 (see Metzger 1998). Tanimuca/Retuara have prefixes that are 'cliticized subject pronouns', and also one 'non-human object pronoun' (Strom 1992: 5). Koreguaje has two prefixes that occur on verbs. The first is ape- 'just beginning to'. The second is ka- and is more difficult to define. When ka- is prefixed to stative verbs, it indicates the process involved in arriving at that state; and, when it is prefixed to active verbs, it indicates an additional and similar action (Cook and Criswell 1993: 73-4 plus p.c.).

3.1 The verb

The agglutinative tendencies of the Tucano languages are best exemplified in verbs. Independent verbs are minimally composed of a verb root plus an ending which indicates person, number, tense and evidential category. Between the stem and the ending, a number of aspect and modality morphemes may occur, as well as miscellaneous suffixes indicating passive voice, causation and motion. Verb stems may be transitive, active intransitive, stative intransitive, negative, quotative or qualitative. Most of the languages have a suffix that distinguishes recent from distant past.

Serial verb constructions are common, and often the meaning of the construction is evident from the meanings of the stems. Note the following examples: 'to hit' plus 'to find' is 'embrace, patting on the back'; 'to hit' plus 'to take' is 'to grab, or snatch'; 'to hit' plus 'to fold' is 'to split'; etc. However, some compounds are not quite so transparent, so that while 'to walk' plus 'to know/ability' is 'to be able to walk', the compound 'to hear' plus 'to know/ability' unexpectedly results in 'to understand'.

3.1.1 Evidentials

The evidential morphemes indicate either how speakers received their information or how closely they identify with what they relate. The evidentials also include person, number and tense information.

In Tanimuca the three suffixes that give evidential information are optional. These optional suffixes indicate: (1) strictly auditory information, (2) assumed information and (3) secondhand information (Strom 1992: 90-1). The suffix that indicates that the information is strictly auditory is consistently used. The other two are quite optional. Thus, the grammar used by Tanimuca speakers does not necessarily convey how they obtained their information (Strom p.c.).

One of the simplest evidential systems is found in Koreguaje. If speakers were present for the event or state they are speaking of, there are no special markers. If they obtained the information from another source, they include an auxiliary verb which indicates that they are not giving a first person account. If they assume that the assertion is/was true, they use an auxiliary verb that indicates probability (Cook and Criswell 1993: 86-7).

Tuyuca is one of the systems that has five evidentials. These evidentials are realized as verb suffixes, some of which are single syllables, and contain the following

1 Gomez-Imbert (1995) also analyses Barasano as having a pitch-accent system.
information: person, number, gender, tense, and evidentiality. The five suffix sets convey the following evidential information: (1) the speaker saw the event or state, (2) the speaker heard, tasted, smelt or felt the event or state, (3) the speaker saw the results of the event or state, (4) the speaker heard about the event or state from someone else and (5) the speaker assumes that the event or state occurred (Barnes 1984: 257). The fourth evidential (secondhand information) only occurs in the past tense, as the speaker necessarily heard the information prior to relating it. The recent past morpheme is inserted when the information was just heard, or was heard recently. For example, if a Tuyuca girl uses an intermediary to express her wishes, the intermediary will use the fourth evidential to say 'She wants to look at that magazine', and it will literally be, 'It was reported to me that she wants to look at that magazine.' The fifth evidential (assumed information) is used when there is no reason to assume that an event did not occur, or is not occurring. If speakers really have no idea as to whether or not an event occurred, they will use the assumed evidential, and will preface their statements with the word /ôba/, which indicates that they are not at all sure. The future tense does not employ evidentials.

Malone (1988) has demonstrated that person, number and evidential information were represented by separate morphemes in proto-Tuyuca. In some of the languages there still exists a string of morphemes, while in others the string has been collapsed into a single syllable.

In the Western north languages, tense suffixes distinguish between masculine singular, feminine singular, and plural. In the Western south language, tense suffixes distinguish between third masculine singular, third feminine singular, first and second singular, and first and second plural. In some conjugations, singular and plural in first and second persons is represented by the same suffix. The Central languages distinguish between masculine singular, feminine singular, animate plural, and inanimate in their person/number markers.

In the Eastern languages, with the exception of Piratapuya and Wanano, the third person evidentials distinguish among masculine singular, feminine singular, and plural. (Note the same distinction in third person pronouns in table 7.5.) In many of the Eastern languages, first and second persons, singular and plural, plus inanimate are all represented by the same evidential, while in others there are person and number markers accompanying the evidentials in non-third-person verbs. See table 7.4 (from West 1980: 28) for a typical example.

3.1.2 Aspect

In most of the languages, aspect is indicated by means of suffixes to the verb stem, and may include: iterative, completive, anticipatory and inchoative.

Table 7.4 Tucano (E2) distant past visual evidentials

<table>
<thead>
<tr>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/2/INAN</td>
<td>-wi</td>
</tr>
<tr>
<td>MASC</td>
<td>-wí</td>
</tr>
<tr>
<td>3 FEM</td>
<td>-wò</td>
</tr>
</tbody>
</table>

Iterative in Barasano

(7) (Jones and Jones 1991: 101)
gahc-ribi bota-ri kēa-kudi-ka-bā īdā
goheri day post PL chop-ITERATIVE-far past VISUAL 3PL

'gahiribi botari kēakudikabā īdā
'The next day they went from place to place chopping down posts
(for the new house).'

Compleitive in Tucano

(8) (West 1980: 56)
di-toha-ʔa
say-COMPLE-EVID:VISUAL

di tohaʔa
'I already said (it).'

Anticipatory in Desano

(9) (Miller forthcoming: § 5.5)
ō-ge āri-ju-ke
here-LOC be-ANTICIPATORY-IMPER

ōge ārijuke
'Stay here in anticipation (of our arrival)'

Inchoative in Koreguaje

(10) (Cook and Criswell 1993: 66)
wai-piʔra-sō aso-ʔmē
hit-INCHOATIVE-INTENSIFIER REPORTED PL

waipiʔrasō asomē
'They began to hit (them).'

7 Tucano
The progressive and perfect aspects are generally achieved through verb compounding. The main verb is in a participial form, and is followed by an auxiliary verb with a tense/person/evidential suffix.

3.1.3 Mood and modality
Imperative and interrogative markers replace the evidential endings, coming at the end of the verb.

**Imperative in Barasano**

(11) (Jones and Jones 1991: 76)
ji-re goti-ja bī
1SG-SPEC tell-PRES.IMPER 2SG

jire gotija bī
‘Tell me!’

**Interrogative in Tucano**

(12) (West 1980: 40)
di-ri
be-PAST.INTERROG

dirī
‘Were you (there)?’ (used for all persons)

Other modality indicators are suffixes which follow the verb root and precede an evidential, imperative or interrogative ending. These mood indicators include: negative, probability/conditional, contraexpectation, desiderative, ability and emphatic.

**Negative and probability/conditional in Tuyuca**

(13) jābika okō peā-ri-atā
yesterday water fall-NEG-DEPENDENT.CLAUSE
bā-ja-wi ūrī-bo-a-hi-ju
2PL-POSS=CL:BUILD burn-COND-RECENT-EVID:PAST.ASSUMED

jābika okō peāriata, bājawi ūrīboahiju
‘If it had not rained yesterday, your house probably would have burned.’

**Contraexpectation and desiderative in Carapana**

(14) (Metzger 1981: 83)
ji u-koa-ga-bī-wi
1SG smoke-COMPL-DESID-CONTRAEXPECTATION-EVID:PAST.VISUAL

ji ukougbīwi
‘I wanted to smoke (but it was not possible).’

### Table 7.5 Tucano (E2) pronouns

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXCL</td>
<td>jiří</td>
<td>išá</td>
</tr>
<tr>
<td>INCL</td>
<td>bří</td>
<td>bří</td>
</tr>
<tr>
<td>MASC</td>
<td>kí</td>
<td>dák</td>
</tr>
<tr>
<td>FEM</td>
<td>kó</td>
<td></td>
</tr>
</tbody>
</table>

**Ability in Tuyuca**

(15) hoá-bāsī-ri-ga
write-ABILITY-NEG-EVID:PRES.NONVISUAL

hoabāsārga
‘I do not know how to write.’

**Emphatic in Yuruti**

(16) (Kinch and de Kinch 1992: 40)
kui-eri-bōkā-ja
be.afraid-NEG-EMPHATIC-IMPER

kūieribōkāja
‘Don’t be afraid!’

3.2 Pronouns

3.2.1 Personal pronouns
The singular pronouns distinguish between first person, second person, third person masculine and third person feminine. The plural pronouns distinguish between first person exclusive, first person inclusive, second person and third person. The same set of pronouns is used for both subject and object pronouns. The Tucano pronouns presented in table 7.5 (West 1980: 177–8) are phonologically representative of present-day forms.

**Tuyuca**

(17) jiří kří-re ījá-wi
1SG 3SG.MASC-SPEC see-EVID:PAST.VISUAL

jiří kříře ījáwi
‘I saw him.’
3.2.2 Possessive pronouns
In six of the languages: C2, El and W1–4, subject pronouns are used as possessive pronouns. The rest of the languages suffix a genitive, generally -ja 'singular' or -je 'plural', to the possessive pronoun and then the item possessed follows. The item possessed is expressed by a classifier or a noun. However, if the noun is a kinship term, the genitive is not used (see Jones and Jones 1991: 61–3 for a description of a typical Tucano possessive system).

Koreguaje
(18) (Cook and Criswell 1993: 45)
\[
\text{dji}³\text{i} \ w\text{a}³\text{thi} \\
\text{1SG} \ \text{knife}
\]
\[
\text{dji}³\text{i} \ w\text{a}³\text{thi} \\
\text{‘my knife’}
\]

Carapano
(19) (Metzger 1981: 184)
\[
\text{kö}³\text{ja}³\text{wi} \\
\text{3SG.FEM-GEN-CL:CYLINDRICAL}
\]
\[
\text{köjawi} \\
\text{‘her cylindrical object’}
\]

Barasano
(20) (Jones and Jones 1991: 61)
\[
\text{i} \quad \text{jiki}³ \\
\text{3SG.MASC grandfather}
\]
\[
\text{i} \quad \text{jiki}³ \\
\text{‘his grandfather’}
\]

3.3 Classifiers
The Tucano languages are primarily numeral classifier languages. Classifiers in the Tucano languages are suffixes that are subdivided into a small set of animate classifiers (which also function as nominalizers) and a larger set of inanimate classifiers. Each classifier denotes salient characteristics of the noun(s) associated with it. Nouns which are not associated with any classifier function as self-classifier suffixes ('repeaters', see Aikhenvald 1994: 420–2), for example, names for some body parts. In all of the Eastern languages plus the Central language Cubeo, classifiers are found suffixed to numerals, nouns, demonstrative adjectives, possessives and verbs (to form nominalized verbs). The other Central language Tanimuca does not suffix classifiers to possessives. In Koreguaje, classifiers are found suffixed to the same categories as the Eastern languages. Full information is not available on the rest of the Western languages. One of the differences between the Western and the Eastern languages regarding classifiers is that in the Western languages nouns with a specific referent are obligatory suffixed by classifiers, although nouns that denote a general category are not suffixed.

Of the 4 Western Tucano languages, Secoya has 17 classifiers, Siona has 20 and Koreguaje has 28. No information is available on Orejón.

Of the 2 Central Tucano languages, Tanimuca has at least 21 classifiers. Cubeo has around 100. The classifiers denote form or function (Strom 1992: 11, and Salser et al.: 16).

Eastern Tucano languages have anywhere from 50 to 140 classifiers. The salient characteristics denoted by these classifiers include not only the three dimensions, but also collection and arrangement. There are rarely heard classifiers, such as one denoting bark that does not cling closely to the tree. By extension, this classifier is used for such things as biggy pants and plywood that has become wet so that the sheets have separated one from another. (For listings of classifier categories, see Barnes 1990 and Jones and Jones 1991: 50–6.)

Classifiers in Secoya
(21) (Johnson and Levinsohn 1990: 41)
\[
\text{mia} \quad \text{sök}²\text{i}³\text{pi} \quad \text{jö}³\text{wi} \\
\text{cedar tree-CL:WITH ROOTS-INST canoe-CL:CONTAINER}
\]
\[
\text{mia} \quad \text{sökijipijow}³ \\
\text{‘a cedar canoe’}
\]

Classifiers in Carapano
(22) (Metzger 1981: 154)
\[
\text{idë}³\text{ji}³\text{ri} \\
\text{chonta-CL:Palm-PL SPECIFIC.NOMINAL.REFERENT-high-NMLZER-CL:Palm-PL}
\]
\[
\text{idëjöri ka}³\text{fö}³\text{ri}³\text{ri} \\
\text{‘chonta palms, specifically the tall ones’}
\]

3.4 Specificity marker
The specificity marker may occur on any nonverb word or construction, and in most of the languages it is obligatory on specific direct objects, experiencers and beneficiaries. In some of the other languages, it marks specific items in the discourse, even
occurring on subjects in Tanimuca (Strom 1992: 7–8). It also occurs on location words (and occasionally on time words) to indicate that the location being referred to will have further significance in the discourse. (See Cook and Levinsohn 1985: 104–8 for a description of a typical Tucano specificity marker.) In all of the Eastern languages, plus C1 and W2, the specificity marker is the suffix -re. In the remaining languages its shape is -Ce, where C is an alveolar consonant.

Objects of the verb which refer to non-specific entities, such as ‘cattle’ (as opposed to ‘that cow’), or ‘a wife’ (as opposed to ‘his wife’), are not marked with the specificity marker in any of the languages. In some instances, this results in noun incorporation. Consider the following examples:

**Tuyuca**

(23) **bādī-kiti**  
husband-have  
‘have a husband’

(24) **di̱kā-kiti**  
fruit-have  
‘have (bear) fruit’

(25) **kape-bādī**  
eyes-not.have  
‘be blind (lit. ‘not have eyes’)’

**Yurutí**

(24) **kīhā bi̱ere kīhā-re**  
3PL~things~SPEC~3PL~SPEC~give-EVID:PAST:VISUAL  
‘They gave the things to them.’

The locative in all but four of the Eastern languages is -pi, E11 and E12 substitute /l/ where the other Eastern languages have /pl/, so that their locative is -pl. The locative in E6 and E7 is -ge. The Central and Western languages have between one and three locatives, which are not interchangeable. In the following example, note that the locative may co-occur with the specificity marker.

**Yurutí**

(25) (Kinch and de Kinch 1992: 13)  
wí-pi-re df wi  
house-LOC~SPEC~belive-EVID:PAST:VISUAL  
wípire df wi  
‘We lived in the longhouse.’

### 3.5 Nouns

The plural suffix for inanimate nouns in the Eastern languages is -ri, and in the Central and Western languages it is -a or -dā. Number–gender suffixes for animate nouns are: -C1 for masculine singular, -(C2) for feminine singular and, in general, -al-dā, -ral-dā, or -na for plural, although there are a number of irregular plural suffixes for animate nouns. C2 does not distinguish between singular and plural for non-human/inanimate nouns (Strom p.c.). The shape of the number–gender suffixes is related to the personal pronouns. There is a set of animate nouns in which the members tend to occur in groups, such as gns and bees. In the Eastern languages the base form of these nouns is plural. To refer to just one gn or bee, a singularizer suffix is needed. That suffix in some of the languages is -bi and in others it is -wi. Inanimate nouns which refer to entities that may occur singularly or in bunches, such as potatoes, firewood, needles, etc., have a general form that does not indicate quantity. To indicate ‘one potato’ or ‘several needles’, either the singular or the plural classifier appropriate to that noun is suffixed to the general form. When indicating a specific quantity (generally one to five items), the plural of the noun or classifier is not used until referring to four or more items.

The locative in all but four of the Eastern languages is -pi, E11 and E12 substitute /l/ where the other Eastern languages have /pl/, so that their locative is -pl. The locative in E6 and E7 is -ge. The Central and Western languages have between one and three locatives, which are not interchangeable. In the following example, note that the locative may co-occur with the specificity marker.

The class of adjectives in Tucano languages includes such categories as demonstrative adjectives, numerals and some quantifiers. Except in C2 which has descriptive adjectives, descriptive terms are generally stative verbs, though some descriptive terms, such as ‘old’ and ‘young’, take the same suffixes as nouns. For example, in Tuyuca, the following descriptive terms are verbs: good, bad, heavy, wide, tall, deep, smooth, sharp, big, difficult, etc., and the colours – dark, light, nature (green-blue),

### 3.6 Adjectives
Janet Barnes

contrast (red-orange). These verbs do not take the full range of verb suffixes, but do function as verbs, as the following examples illustrate.

**Tuyuca**

(26a) dikí-ri-ku
be.heavy-NEG-EVID:PRES.ASSUMED
dikíriku
'It is not heavy.'

(26b) basoki debó-ro ibiá-ki
person be.bigger-than-ADVZ be.tall-EVLD:PRbS.ASSUMED
basok¡ debor6 ibilikí
'He (a gorilla) is taller than a person.'

**3.7 Negation**

Verbs are negated by means of a negative suffix which precedes the evidential (or person—tense) ending. Nouns can be negated by suffixing a different negative morpheme to them. There are generally two negative verbs: 'to not have' and 'to not be'. Concepts such as 'never', 'no one' and 'nothing' are encoded without the use of specific words for these concepts (Barnes 1994: 337–9). Neither is there a negative response word. Rather, the negative response to 'Will you go?' is a full clause 'I will not go.'

**3.8 Nominalized verbs**

Nominalized verbs in the Tucano languages are used where other languages use relative clauses. Note the following examples which are inflected for tense and mood:

**Barasalom**

(27) (Jones and Jones 1991: 43)
bue-go study-PAST-SG.FEM
buego
'she who studies'

**7 Tucano**

bue-ka-ko study-PAST-SG.FEM 'she who studied'
bue-ro-ko study-FUT-SG.FEM 'she who will study'

**3.9 Switch-reference**

In the Eastern languages, if the subject of a dependent clause is the same as the subject of the clause following it, there is a set of suffixes, often indicating person and number, that is used on the verb of the dependent clause. However, there is only one suffix to indicate that the subject of the dependent clause is different from the subject of the following clause. The subject of that clause is the same as the subject of the following clause. This type of construction is typically used when describing a series of events that took place one right after another. (For a discussion of a typical Eastern Tucano switch-reference system, see Longacre 1983.)

**Different subject in Wanano**

(28) (Longacre 1983: 202)
to wa?atí tidá thuataşi
tidá thuataşi
'This when.going-DS they won't{return}.
'to wa?atí tidá thuataşi
'Then goes, they won't return.'

**Same subject in Wanano**

(29) (Longacre 1983: 202)
tiro wa?aro thuataşi
3MASC when.going-SS won't{return}
tiro wa?aro thuataşi
'When he goes, he won't return.'

**Different subject in Tuyuca**

(30) bií wáa-ri wáa-ida
2SG go-DS go-SG.MASC.FUT
bií wáari, wáaída
'If you go, I will go.'

---

5. The Western language Koreguaje and the Central languages do have a negative response word. In Koreguaje it is pláaní. The negative response words in the Central languages are similar to negative response words in one or more Arawak languages. In Cubeo it is bi, and in Tanimuca/Retuarã it is berä?ô.

6. Tanimuca/Retuarã has both relative clauses and nominalized verbs.
The Western north languages utilize same-subject and different-subject markers between sentences. Typically, the verb of the preceding sentence is repeated, and suffix(es) are added to indicate same subject or different subject. (For examples, see Wheeler 1987: 179–80.)

4 Syntax

The Tucano languages are typically AOY, SY languages, with variations in word order due to discourse considerations. In some of the languages, if the actor is an animate noun, whether expressed or understood, the corresponding pronoun is mandatory just prior to the verb. Thus, these languages have a large percentage of sentences that are OAV, with the A argument being a pronoun. The two Eastern south languages, Barasano and Macuna, are OV7 and OVA respectively, and the Western language Koreguaje is VAO.

Typically the object is unexpressed if understood, so that it is common to hear a command such as 'Close!', rather than 'Close it!' Also, statements such as: 'I already told' rather than 'I already told it to her' are common. When they do occur, the direct object precedes the indirect object (see example 24).

In some of the Eastern languages, such as Tuyuca, utterances generally involve many short sentences, with a marked lack of hypotaxis. Others, such as Taiwano, tend to have sentences with long series of dependent clauses.

4.1 Time and location

There is flexibility within the sentence regarding the order of words or phrases denoting time and location. In general, expressions of time precede the verb and may follow a subject pronoun. Expressions of location generally follow expressions of time, and often follow the verb. Expressions of location may also signal paragraph boundaries (Waltz 1975: 50).

4.2 Noun phrase

Noun phrases are not commonly used in most Tucano languages. If, in a given sentence, the context makes it clear that a canoe is being discussed, then to state that it is big requires only that the classifier referring to canoes be suffixed to the word 'big'. Phrases such as 'My big, old, beat-up canoe' are not often heard in most of the languages. Generally just one descriptive word will occur in a given sentence. But if all of the descriptive concepts are new to the context, and necessary for full understanding, then each descriptive term will occur with the appropriate classifier attached to it. The following example is taken from a text about a DC-3 aeroplane that landed upriver. The 'hollow' classifier, which refers to the aeroplane, is also that used for canoes and reed-like plants which grow with hollow centres.

Tuyuca

(32) tií-wí pai-ri-wí-bêdá
that-CL:HOLLOW big-SG.NMLZR-CL:HOLLOW-INST
weki-a pi-a rÀ tikoko-wa
cattle-PL two-PL.ANIM send-EVID:PAST.VISUAL

'Ve two bulls were sent in (by means of) that big plane.'

BIBLIOGRAPHY


Languages of the Pano family are spoken on the eastern side of the Andes in Peru and nearby regions of Brazil and Bolivia. They were first recognized to constitute a family, on the basis of the comparison of seven languages, by De la Grasserie (1890). He named it 'la famille linguistique Pano' after one of the languages, called Pano or Waripana. The languages show close similarities, indicating a fairly shallow time-depth and recent expansion and split.

It has been suggested that the Pano family may be genetically related to the Tacana family, which is discussed in chapter 13 (there is phonological comparison and reconstruction in Key 1968 and Girard 1971). There are some lexical and some grammatical similarities, e.g. 2sg pronoun mi- and the forms of switch-reference markers. However, more work is required to determine whether these similarities are sure indicators of genetic connection, rather than being due to areal diffusion.

Shell (1965/1975) undertook a systematic comparison and reconstruction of proto-Pano phonology and aspects of the morphology. She listed more than eighty 'language' names found in the literature but some of these are names of dialects or clans within a language group, or alternative names of languages. The number of distinct languages is probably no more than thirty, as shown in table 8.1; some of these may turn out, on further study, to be dialects of a single language. Many different subgroupings have been suggested in the literature. In table 8.1, I recognize three subgroups and a number of ungrouped languages; the subgroupings are made, tentatively, on the basis of shared phonological and morphological characteristics and— to a limited extent— shared vocabulary. Note that several of the names end in -nawa 'people' and some others in the plural marker -bo. Kensinger (1985) provides a useful overview of Pano groups with bibliography.

1 I owe special thanks to Ivagene Shive for Yaminawa, Eugene and Marie Scott for Sharanawa, Kim Fowler for Yoranawa, Harriet Fields for Matses, and Wayne Gill of NTM-Bolivia for Chimane.
One often finds many lexical roots shared between two languages, although the forms of the suffixes vary, leading to initial lack of intelligibility. However, this can be overcome through a degree of exposure to another language. After a few days of adjustments to differences, a Yaminawa or a Sharanawa can understand most of a give-and-take exchange with a Yoranawa but gets lost in flowing discourse. The Sharanawa are reported to be able to understand much of the Yawanawa. When Cashinawas converse with Sharanawas they are known to resort to a kind of Pano by suppressing the use of most suffixes, especially those not shared between them.

Note that, unless otherwise indicated, all examples in this chapter are from Capanawa, the language with which I am most familiar. Also, allomorphic forms are used in examples without identifying the basic form of the morpheme each time.

---

Table 8.1 Languages of the Pano family

<table>
<thead>
<tr>
<th>No.</th>
<th>Language</th>
<th>Speakers</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yaminawa</td>
<td>500 P</td>
<td>Br</td>
</tr>
<tr>
<td>2</td>
<td>Amawaca</td>
<td>200 P</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Cashinawa/Honikoin</td>
<td>500 P</td>
<td>Br</td>
</tr>
<tr>
<td>4</td>
<td>Sharanawa/Shanindawa/Chanindawa/Inonawa/Marinawa</td>
<td>300 P</td>
<td>Br</td>
</tr>
<tr>
<td>5</td>
<td>Yawanawa</td>
<td>200 Br</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Chitonawa</td>
<td>35 P</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Yoranawa/Nawa/Parquenawa</td>
<td>200 P</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Motonawa</td>
<td>300 Br</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Mastanawa</td>
<td>100 P</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Chacobo</td>
<td>400 Bo</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Arazaíre</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Atsawaca</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Yamiska</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Katikina/Camannawa/Waninawa</td>
<td>300 Br</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Pacawara</td>
<td>12 Bo</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Capanawa/Pahenbakebo</td>
<td>400 P</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Shipibo/Conibo/Xetebo</td>
<td>8,000 P</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Remo</td>
<td>Br</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Marebo</td>
<td>400 Br</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Wariapano/Panono/Pano</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Isonoawa</td>
<td>30 P</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Canamari/Taveri/Matoihal</td>
<td>Br</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Cashibo/Cacataibo/Comabo</td>
<td>100 P</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Culino Br</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Karipuná Br</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Kaxariri</td>
<td>100 Br</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Mattes/Mayuruna</td>
<td>2,000 P, Br</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Nokarnán Br</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Poyanawa Br</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Tutxinawa Br</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:
The approximate number of speakers is given for each language; † indicates a language that is probably no longer spoken. Locations are indicated by P, for Peru; Br, for Brazil; Bo, for Bolivia. (For approximate geographical locations of each language see map 7.) Note that in Spanish 'hu' would be substituted for 'w' in these names.
Table 8.2 Proto-Pano consonants

<table>
<thead>
<tr>
<th>voiceless stop</th>
<th>apico-</th>
<th>apico-</th>
<th>apico-palatal</th>
<th>dorso-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilabial</td>
<td>alveolar</td>
<td>palatal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless affricate</td>
<td>t</td>
<td>ts</td>
<td>tf</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>voiceless continuant</td>
<td>s</td>
<td>f</td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>voiced continuant</td>
<td>β</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>semivowel</td>
<td>w</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2 PHONOLOGY

2.1 Segmental phonology

There are four proto-Pano vowel phonemes: low a, high front unrounded i, high back unrounded õ and high back open rounded o. All have been retained with a high degree of consistency. Matses now also has a mid unrounded vowel e; in Cashibo the ai sequence has developed into e and the sequence aw becomes o.

Of the sixteen proto-Pano consonants, the stops and affricates have most consistent reflexes in daughter languages.

Though the voiced bilabial fricative ß is not a resonant in phonetic terms, it is here classed with r because it patterns like r in not being able to serve as the first member of a consonant cluster and like r it blocks some syllable reduction rules. Possibly ß was not a fricative in proto-Pano; it is realized as a voiceless bilabial spirant in the languages of the Yaminawa group.

The glottals have been lost in most daughter languages. Both ß and h are retained in Amawaca, Chacobo, Capanawa and Pacanawa, just h in Yawanawa, Shipibo and Wariapano, and just ß in Camanawa.

The reflexes of the nasal consonants range from nasals to oral plosives in the Yaminawa group.

Sharanawa and Cashinawa have lost the distinction between s, f, ẓ, ñ in surface structure m and n do not occur in syllable-final position and hence do not serve as the first member of a consonant cluster except in Matses. However, nasalized vowels preceding obstruents typically show a transitional phonetic closure which produces an apparent epenthetic [n] or [m] which some investigators have taken as a neutralization of n or m. Examples like the following show the seeming displacement because of the pre-consonantal closure of the oral passage:

(i) Nasal spread

Syllable-final m and n reduce to leftward-spreading nasalization on any preceding contiguous string of vowels and semivowels, and when the nasal consonant is followed by a glide (w, y, h, ß?) the nasalization also spreads rightwards onto contiguous vocalic and semi-vocalic segments. This nasalization rule is missing in Chacobo and to a large extent in Yaminawa, Sharanawa and Yoranawa where m and n tend to be denasalized, becoming plosives [b] and [d] except when there is a following nasal vowel in the word.

Yaminawa

(1a) [ado] (b) [anã]
ano
majãs-ERG
'majãs' (edible rodent, O form) 'majãs' (edible rodent, A form)
(2a) Oral vowels: (b) Leftward spread: (c) Rightward spread:

hawina [hãwĩ goũo] [wirãwĩ] hawin goũo
wiran-wi
push-IMPER
'his (own)' 'his house' 'Push it over!'

Unlike the continuants s, f, ẓ, ñ in surface structure m and n do not occur in syllable-final position and hence do not serve as the first member of a consonant cluster except in Matses. However, nasalized vowels preceding obstruents typically show a transitional phonetic closure which produces an apparent epenthetic [n] or [m] which some investigators have taken as a neutralization of n or m. Examples like the following show the seeming displacement because of the pre-consonantal closure of the oral passage:

(ii) Vowels are elided in a weak position such as in a stem-final unaccented open third syllable.

Yamarinawa

(3a) [hoʃũŋkiʔ?] (b) [ku-tãmãʔ?] hoʃũŋ-ha-kin ku-tan-pan
red-make-CONJ go-DIST-shall
'makes it red' 'I shall go away.'

(4a) [nonomã] (b) [nõo] nonoma-n nonom
duck-ERO
'duck' (A form) 'duck' (O form)
(iii) Syllable-final stops p, t, k, m, n are elided; note that m and n function as stops but b and r function as resonants.

Chacobo
(5) kapiti > kapit > [kapi]
'alligator' (O form)

(iv) h is elided word-internally.
The phoneme h occurs only morpheme-initially but morpheme concatenation can make it follow another segment word-internally in the underlying form (where it can then be the cause of the deletion of a preceding nasal consonant before it is subsequently suppressed).

(6a) [himiiwi] (b) [haini]
himi-hiwi ha-hini
blood-tree mouth-water
‘bloodwood’ ‘saliva’

(v) Vowel harmony
In some contexts an a vowel fully assimilates to the vowel of the preceding syllable.

We get assimilation in (7b) but not in (7a).

(7a) [tiʔkaki] (b) [tiʔkiti]
tiʔka-kin tiʔka-t-i
break-did break-REFL-PRES
‘He broke it.’ ‘It breaks.’

In the Yaminawa group, i assimilates to an i vowel of the following syllable.

Yaminawa
(8) [pitsiki] (b) [piʔki]
pi-tiski
shoulder-push
‘to push on the shoulder’

(vi) Alternate mora timing
An odd-even syllable-timing characteristic common in Pano languages causes phonological modifications such as segment deletion, plosive nasal release, stress assignment and possibly vowel harmony.

In Shipibo the morpheme -riʔba- changes to -riʔbi- when the syllable ña is even-numbered from the beginning of the word:

Shipibo
(9a) ñina-riʔbi-wi (b) ka-riʔba-i
look-again-IMPER go-again-PRES
‘Look for it again!’ ‘I am going again.’
— A syllable-final glottal is elided in even-numbered syllables, as in (10b):

(10a) ñina-riʔbi-wi (b) biʔ-riʔbi-wi
look-again-IMPER take-again-IMPER
‘Look for it again!’ ‘Take it again!’
— Stress assignment often depends on the syllable structure of the word, but details vary from language to language. In Capanawa, Shipibo, Canamari and some others of the Capanawa subgroup, a closed second syllable is stressed. This stress assignment takes place before syllable-final consonant deletion so that some surface open second syllables remain stressed:

(11a) [jisis] (b) [bari]
hisisi bari-n
‘a variety of ant’ ‘the sun’ (A form)
— An intervocalic nasal consonant before an oral vowel takes on a plosive release when it begins an even-numbered syllable:

Amawaca
(12a) [hondi] (b) [honí]
honi honi-n
‘man’ (O form) ‘man’ (A form)
— A final odd-numbered syllable of certain verb affixes is elided, e.g. the suffix -paki- ‘all’ retains its final syllable if that syllable is even-numbered:

Sharanawa
(13) [yoipakitiroki] yoi-paki-tiro-kin
tell-ALL-well-having
‘having told it all well’
— The final syllable of the suffix is lost when the syllable is odd-numbered:
3 GRAMMAR

3.1 Word classes and noun phrase formation

The open word classes are Nouns, Verbs, Adjectives and Adverbs; Adverbs are mostly limited to time and location words. Pronouns, Demonstratives, Numbers, Postpositions and postpositional Particles form closed sets. Nouns and Adjectives are easily verbalized by adding verb suffixes; if no transitivizing suffix is present the meaning is 'to be...' or 'to become...', otherwise the meaning will be 'to make or cause...'. Typically, if the verb is in present tense and is not needed to carry verbal suffixes, both verb and tense may be omitted.

(14) [opaki] o-paki-kin 'having come down'

3.1.1 Noun phrase

The noun phrase can consist of a head and a series of modifiers. There are no definite or indefinite articles; nouns may be used in the adjective modifier slot:

Quantifier - Possessive - Adjective - Head-noun - Adjective - Numeral

For example:

(19) hawin şibon ka’iti pono romifin

his palm leaf back veins small

'the little veins of the back side of his palm leaves'

Possessives are formed by appending a possessive suffix to a noun or pronoun.

Some adjectives are by preference preposed to the head, others are postposed, and some may occur in either position. When preposed, adjectives are subject to the phonological rules that shorten them. Preposed adjectives have a delimiting function.

(20a) hoa paJini (b) paJi hoa

flower red red flower

'the reddish flower' 'a flower that is red'

Two nouns may be compounded, in which case the combination is subject to the shortening rules.

(21) [maribí] mari-bina aguti-wasp

'a black wasp' (a wasp that is black like an aguti)

3.1.2 Pronoun classes

Pronouns occur in free form for focused A and S, non-focused A and S, and O functions. Bound forms are used when an oblique case-marking suffix such as indirect object is added. Not all Pano languages show the focus distinction but it is found in each of the major groups.

(22) ?ian 'I' (free form)

?i?-ki 'me-to' (bound form)
3.1.3 Relative clauses
The head noun is shifted out of a relative clause or omitted altogether.

(23) ?ani hiwi mebi teșpat-ai teșpa hanin ?iso honit-i big tree branch bifurcate-PRES fork LOC monkey hide-PRES
'A monkey is hiding in the fork of a branch that bifurcates from a large tree.'

3.1.4 Noun phrase pluralization
Plural is usually marked only on the verb rather than on an NP in A or S function; the plural on NPs in other functions is marked only when it is essential for disambiguation. The plural marker for nominals is -bo, homophonous in many instances with the generic marker -bo: honi-bo (man-pl) 'men' or (man GENERIC) 'mankind'. (The suffixes and postpositions are appended to the NP, not just to the head noun.)

(24) ?oá tsaʔot-ai-bo his-i sit-PRES-pl see-IMPER
'Look at those (who are) sitting over there!'

3.2 Transitivity concord
All Pano languages are characterized by a distinctive transitivity concord system. Transitivity concord marking is interrelated with the systems for identifying verb transitivity, switch-reference, sequential action, adverbs, relative clauses, temporal clauses and locative expressions. The usual conditions are:

(a) Transitivity is semantically determined: a verb phrase is transitive when it has an O referent that is different from the A referent. The referent-identifying noun or pronoun need not be overtly expressed.

(25a)  haa-ta his-i-ki 3-A DECL see-PRES-FACT
'He sees [it].'

(25b)  haa ta his-it-iʔ-ki 3S DECL see-REFL-PRES-FACT
'He sees himself (in a mirror).'

(b) Pano languages show a variety of split systems, the marking of the A and S being affected by focus.

(26a)  maya-paki-wi turn-fell-IMPER
'Knock it down spinning!'

(26b)  mayápakitiştikí turn-REFL-fell-REFL-PAST-3-FACT
'He fell spinning.'

3.2.1 Switch-reference
Pano switch-reference is marked by suffixes on subordinate verbs that also indicate the time of the subordinate verb relative to the main verb. A subordinate verb may precede or follow its main verb. Table 8.3 shows the Capanawa suffixes. There are corresponding forms in most other Pano languages. (Shipibo forms are given in Loriot, Lauriault and Day 1993: 54-7.)

The following sentences exemplify these suffixes:

(27) hiwi ?ani baʔo kiyani bimi tfaʔot-aš naman tree big baʔo tall fruit soft become-SWa below
mi-rakaʔ-tai-ton ska honon mira-šon pi-kin earth-ON-lie-PRES-SWc then peccary find-SWb eat-CONJ
'The baʔo tree is huge, very tall. When the fruit becomes soft and falls to the ground the peccary finds it and eats it.'
Table 8.3 Switch-reference markers in Capanawa

<table>
<thead>
<tr>
<th>verbal suffix</th>
<th>subordinate clause</th>
<th>main clause</th>
<th>function of coreferential argument in</th>
<th>action of subordinate clause in relation to action of main clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWa -a3</td>
<td>S</td>
<td>S</td>
<td>before or same time</td>
<td></td>
</tr>
<tr>
<td>SWb -yon</td>
<td>A/S</td>
<td>A</td>
<td>before or same time</td>
<td></td>
</tr>
<tr>
<td>SWc -ton</td>
<td>A/S</td>
<td>O</td>
<td>same time</td>
<td></td>
</tr>
<tr>
<td>SWd -a7</td>
<td>O/IO</td>
<td>A/S</td>
<td>before or same time</td>
<td></td>
</tr>
<tr>
<td>SWe -nogon</td>
<td>A/S</td>
<td>A/S</td>
<td>follows</td>
<td></td>
</tr>
<tr>
<td>SWf -ybo</td>
<td>A/S (must be plural)</td>
<td>O</td>
<td>before or same time</td>
<td></td>
</tr>
</tbody>
</table>

(28) mai kinti ki kobinha-?a bi ki siniska-a3 b?mikii?i-i clay pot in cook-SWa but REP old.get-SWa cracks-PRES 'When anyone cooks in a clay pot that is old, the pot cracks.'

(29) ?ot?i ka?i-?i-?a ho-ri?i-?i-ton b?ilti-?i-?i min far going-PAST-SWa come-AGAIN-PRES-SSwc meet-CONJ you ho-ri?i-?i-ton ya?i-kin hawi-ya hinima-son come-AGAIN-PRES-INTERROG say-CONJ come-SWg rejoice-SWb bi?ilti-?i-?i-?i hinima-son met-CONJ very happy-CONJ 'He had gone a great distance and upon his return she met him with the words "Have you returned?" Because he came, she rejoiced and went out to meet him. So she was very happy.'

(30) nonii ?aniha-?o min pitibo.ri?i-bi banawi hano canoe enlarge-SWb your food also plant there napo-ya?i-no?o pile-FUT-SWe 'After enlarging your canoe, plant your food also in order to pile it in the canoe later.'

(31) yo?i-a kobinha-?o bimakanya mi?i-a pa?a-?i-ton yucca cook-SSwb on the surface of raw-be-SSwc birabi-kin rami?o ska kobin-non turning-over-CONJ other-side then cook-SWh

(32) hatii?i-bi hato yo?i-wi honan-a?bo all them tell-IMPER know-SWf 'Tell all of them so that they will know.'

3.2.2 Adverbial suffix concord

Some adverbial verb suffixes have pairs of selectional variants according to the transitive or intransitive valence of the verb; the adverbial suffix must agree with the transitivity of the verb. For example, in Capanawa the variants for 'abruptly, rapidly, completely' are -ka?in- (intransitive) and -ba?in- (transitive) and the variants for 'movement towards the speaker' are -ki?ran- (intransitive) and -be?lan- (transitive).

(33a) mapi-ka?in-i go.up-rapidly(INTR)-CONJ 'He goes up (the bank) rapidly.'

(33b) bi?i-ba?in-kin grab-rapidly(TR)-CONJ 'He grabs it abruptly.'

(33c) wao?i?-kir?an-i haw-i cry-coming(INTR)-CONJ come-CONJ 'He came crying.'

(33d) his-binan-kin bi mi?i-yama-kin see-coming(TR)-CONJ but touch-NEG-CONJ 'He saw it as he came but didn't touch it.'

3.2.3 Locative phrases

Locative phrases in sentences without a verb and locative phrases that relate to the O take no transitivity marking. Locative phrases relating to the source or position of the A carry transitive marking; those that modify S take intransitive marking.

(34a) hano ta? ki there DECL FACT 'There he is.' (no verb present)

(34b) hano ?a?o his-kin there majas see-CONJ 'He saw the majas (rodent) there.' (the locative modifies the O)
3.2.4 Ergative marking

(i) Form

In most Pano languages, any syllable-final nasal consonant is realized as nasalization on the preceding contiguous vowels and semi-vowels. Hence nasalization is the only clue to the presence of the ergative marker -n, but in some of the languages the nasalization has been lost.

In Chacobo and Amawaca, if the verb is in present tense the A or S is automatically postposed to follow the verb where it is not marked as A, and is then identical to its 'citation form' (i.e. as an answer to 'What is that?') which is not the same as the O form.

Chacobo
(35) inaka 'dog' (citation form)
    ina 'dog' (O form)

Syllable reduction has reduced the word for 'dog' above to two syllables. Switching the order of the A and O makes no difference to the marking:

Chacobo
(36a) kamanó ina pi-ki
    (b) ina kamanó pi-ki
    tiger dog eat-PAST dog tiger eat-PAST
    'The tiger ate the dog.'  'The tiger ate the dog.'

At first glance it appears that the reduction of the final syllable of the O noun is an O-marking device, but it is rather simply pre-posing the noun to a co-constituent that causes the reduction of the final syllable, as shown when the noun is preposed to an adjective or postposition:

Chacobo
(37) ina pistia
    dog small
    'small dog'

(ii) Function

There are a number of kinds of ergative system in Pano languages.

(a) Some, such as Wariapano, have a split system, with nouns having S and O marked in the same way (absolutive case) and A differently (ergative), but pronouns having S and A marked in the same way (nominative case) and O differently (accusative).

(b) Others, such as Shipibo, have both nouns and pronouns showing an absolutive (SO) versus ergative (A) system.

(c) Capanawa has a third type of organization. Here emphatic pronouns are like nouns in having ergative suffix -n marking A function, with S and O left unmarked. Non-emphatic pronouns show a different system with -n marking both S and A functions, while O is left unmarked. The forms of the 1st person pronoun are:

Chacobo
(38a) ké-ki ba7í-no
    go-PRES path-on
    'He is going on the path.'
(38b) ké-ki şobo-no
    go-PRES house-in
    'He is going home.'

Moving the noun with its locative suffix into preverbal position causes the loss of the final consonant and vowel from ba7í-no 'on the path'.

Chacobo
(39) ba7í kó-kiy-a
    path-(on) go-PROG-1sg
    '(I) am going on the path.'

Similarly in Capanawa the locative -no is preserved when attached to single-syllable forms like the demonstrative ha- 'that': ha-no (3-LOC) 'there'; but when the locative is affixed to a noun the reduction rules reduce -no to nasalization: şobo-no becomes şobö. Likewise the instrumental -7an and the possessive -na are reduced to nasalization except when attached to a single-syllable form or when not preposed to the head noun, as in hawai ‘his’.

The Pano languages have a number of types of ergative system:
A clause can include just one instance of a non-emphatic pronoun, as in:

(40a) intransitive
ra?ma ta? ?i-n ka-?i
now DECL 1-SA go-PRES
'I'm going now.'

(40b) transitive
kapí ta? ?i-n his-i
alligator DECL 1-SA see-PRES
'I see an alligator.'

A pronoun can be repeated at the beginning of a clause to show that it is in focus. As long as it is non-emphatic, a nominative (marked by -n)/accusative system is still followed; for example:

(41a) intransitive
?i-n ta? ?i-n ka-?i
1-SA DECL 1-SA go-PRES
'I am going.'

(41b) transitive
?i-n ta? ?i-n pi-?i
1-SA DECL 1-SA eat-PRES
'I am eating [it].'
3.4.2 Suffixes

In some languages more than 130 verb suffixes are available. The suffixes tend to fall into 9 groups, each group having semantic or functional features in common and sharing a preferential order with respect to the other groups. Beginning closest to the verb stem, the general order is (1) thematics, (2) negation, (3) motion, (4) aspects, (5) person and tense, (6) mood and subordinate markers, (7) evidentials, (8) negation, (9) connectives. Each group contains some suffixes that are semantically mutually exclusive with one another and others not mutually exclusive but which follow a preferential order within the group. The further to the right their occurrence in the string of suffixes, the greater the corresponding semantic scope regarding other constituents of the clause.

Group 1. Thematics

Group 1 suffixes complete the thematic structure of the verb, such as morphemes that signal reflexive, reciprocal, causative, transitive, benefactive, detrimental and dative of interest.

A reflexive is always intransitive.

(45) [tamáni?tasJK] ta-má?ni-t-a-š-ki
foot-switch-PREVF-PAST1-3-FACT
‘He took a step.’ (lit. ‘He switched his foot.’)

(46) mayápakitipíki
maya-t-paki-t-ipi-f-ki
turn-PREVF-PAST2-3-FACT
‘He fell spinning.’

(47) haa ta ?a?mi-bi ñí?ti-t-a-š-ki
3S DECL 3-REFL-kill-PREVF-PAST1-3-FACT
‘He killed himself.’

Reciprocals are marked by the thematic suffix -nan and are always intransitive. In a reciprocal clause the S NP refers to the group of participants, for example,

(48) haa-bo rabi ta his-i-nan-i-ki
3-pl two DECL see-RECIP-PRES-FACT
‘Both of them are looking at each other.’

(49) haatian-bi ta his-i-nan-ipi-kana-š-ki
there-INTENSIFIER DECL see-RECIP-PAST2-PL-3-FACT
‘Right there they looked at one another.’

Group 2. Negatives and strong affirmatives

There are two possible locations in the verb suffix string for negatives: group (2) -yama- ‘not’ and group (8) -ma ‘not’. The morpheme -yama- follows the thematics, may be interspersed among groups (3) and (4), and always precedes the tense or terminal group. In Shipibo, Capanawa and many other languages, -yama- is used in all forms of finite and subordinate verbs, but nouns, predicate nominals, nominalized verbs and relative clauses are negated by postponing -ma to the tense or to the nominal. In Yaminawa and Sharanawa -yama- is used only in verbs having a future or incomplete tense or aspect and in the imperative mood, otherwise -ma is used, posted to the tense as described above. In some languages -yama- also functions as a verb stem:

(50) yama ta ha-i-ki
not DECL AUX-PRES-FACT
‘There aren’t any.’ or ‘There is nothing.’

(51) mia ta? ?i in yama-ha-?ipi-ki
you DECL lsg NEG-make-PAST-FACT
‘I missed you.’

Camílnawa

(52) ði-yama-wi
see-NEG-IMPER
‘Don’t look!’

Lleonawa

(53a) ?i?ah ka-i-mba i
3S go-PRES-NEG BE-PRES
‘I’m not going.’

(53b) ka-yamba-wi
go-NEG-IMPER
‘Don’t go!’

Group 3. Adverbs of motion such as upward, downward, circular, to one side, along another referent, inverted, direction to or away from the speaker, action done over a distance.

Group 4. Aspectual adverbs of daytime or night-time activity, quickly, desiderative, truly, intensively, first, just, next, perhaps, again.

Group 5. Person markers, plural markers, tenses, aspects of duration. Groups 5 and 6 are generally final; at least one from either group is generally obligatory on the verb though discourse constraints allow variation.
Group 6. Markers of subordinate verbs, nominalization, conditional, permissive, conjunctive, interrogative and imperative.

The interrogative indicators vary in form and location; some attach to the verb and others are postposed to the initial major constituent of the sentence:

**Shipibo**

(54) mi-a ki ho-a

2-S INTERROG come-PAST

'Have you come?'

**Capanawa**

(55) min haw-i-n

2S come-PRES-INTERROG

'Have you come?'

Group 7. Contains evidentials (e.g. 'factual', 'reported', 'assumed') and also mood markers and other morphemes that indicate intended illocutionary force, and though they attach as suffixes to the verb some may also be fronted to the beginning of the sentence where they attach to whatever is focused. In the answer part of (56) there are two evidentials: -s- 'visible' and -ki 'fact', and the declarative mood marker -ta-.

(56) Question:

mama ka?ri?i?

mother what.about

'Where is (your) mother?'

Answer:

?o?i-s-ta haw-i-ki

there-VISIBLE-DECL come-PRES-FACT

'There she comes, (visibly, in fact).'

Group 8. Mutually exclusive with some of group 7 and contains only the negative -ma.

Group 9. Contains logical connectives of various types, e.g. 'because', 'since', 'when', 'of course', 'indeed', 'but', 'although', 'even though', 'for'. They are more loosely bound to the verb than the other groups and are often found written as independent particles. The Capanawa connective kin 'because' is an example:

(57) haa sho-bo-7o nakat-ai-kj ta? ?i haw-mpi-ki

that house-in lie-PRES-because DECL 1S come-PAST2-FACT

'I came because she was lying in that house (the hospital).'

There may be more than one occurrence of affixes in groups (1)-(4) in the same verb. For example, the causative -ma- introduces an O argument to the verb. When more than one causative occurs, each introduces a new argument. In (58) the arguments are differentiated by subscripts.

(58) bi?i-ma-ma-wi

grasp-CAUS-CAUS-IMPER

'(You) make (him1) get (him2) to get (it).'

(59) min bene kiin-yama-yama-wi

your husband want-NEG-NEG-IMPER

'Do not despise your husband!' (lit. 'Do not not-love your husband!')

There are generally strong preferences in the order of suffixes between groups 1 and 4, and similarly within each group there is a preferential internal order; when the order changes there is a corresponding difference of scope, the rightmost suffix having the widest scope.

(60) his-ma-r?bi-wi

see-CAUS-again-IMPER

'Show it to him again!'

(61) his-ribi-ma-wi

see-again-CAUS-IMPER

'Make him take a second look!'

Within groups 5, 6 and 7 the order is fixed, though in some languages the plural marker precedes certain tenses and follows others. Some Pano languages have four or more past tenses but the future tense set does not always mirror the past tense set in the spans of time covered.

All Pano languages seem to have a large number of verb stems that terminate in -i or -a; -i verbs are intransitive, "be/do"; and -a verbs are transitive, "do". The form of these verbs must be lexicalizations of periphrastic constructions formed on the auxiliary verbs ʔiʔk- 'to be, do, say' and ʔaʔk- 'to do', e.g. Capanawa tofo- 'to burst', po?a- 'to crush'.

Repetition of part of the verb to show duration, intensity or repetition can include some of the suffix string (up to and including group 4) attached to the verb stem:

(62) han pi-pana n ta? min hini-baʔin

3S eat-would DECL 2 release-completely

hini-baʔin-ai

release-completely-PRES
3.5 Mood indicators

Pano mood indicators are part of suffix group 7 and are either (1) indicators of intended illocutionary force, or (2) indicators of speaker attitude towards the addressee or towards the articulation. All modify the whole sentence and tend to occur either at the end of the verb suffix string or postposed to the first major constituent of the sentence.

3.6 Deictics

It is common to find that, in discourse, explicit NP reference is kept to an absolute minimum. The switch-reference system enables listeners to identify many referents, although demonstrative pronouns are sometimes used to point to a referent since there are no definite or indefinite articles. Pronouns may be anaphorically omitted, unless needed for disambiguation.

In general the identification system for the referents of focus and topic is that, after a referent has been put in focus by fronting, it is thereafter referred to by 'zero' pronouns (i.e. pronominal reference to it is omitted) but a topic referent and demonstrative pronouns are sometimes used to point to a referent.

In dialogue the demonstrative pronouns are used to differentiate referents, for example in Capanahua:

(63) niá 'this' (near the speaker)
?óá 'that' (away from speaker)
tóá 'that' (near the addressee)
haa 'that' (whatever has been referred to in context)
nino 'here' (near the speaker)
hano 'there' (away from the speaker) or (the place previously referred to)
?ono 'there' (away from speaker and addressee)
?onu 'over there' (far from speaker and addressee)
The Makú family consists of four languages belonging to seven tribes—see table 9.1. One is spoken around the Middle Rio Negro in Brazil, with the others ranging from the Upper Rio Negro to the region of the Vaupés and Japurá spanning Brazil and Colombia (see map 8). The two groups belong to different linguistic areas and greatly differ in their grammatical structure.

The Makú are nomadic hunters and gatherers. They have a low socio-economic status among other indigenous peoples of the region—speakers of Arawak and Tupano languages, who specialize in slash-and-burn agriculture. The Makú live in small groups along little streams, avoiding rivers, and providing their abodes with fish-poison (timbó) and game, receiving manioc in exchange. As jungle dwellers, they are believed to possess supernatural magical power. Like the Pygmies in Africa, they are despised and feared at the same time.

We are grateful to Elias Coelho for providing information on Hupda.

1 We are grateful to Elias Coelho for providing information on Hupda.

2 Puinave—not spoken in the region of the Inirida river in Colombia (see chapter 13 below)—has sometimes been linked to Makú, as a Makú-Puinave family. In fact, no genetic relationship between Makú languages and Puinave has as yet been proven.

The term ‘Makú’ is used as a pejorative term for Indians in the regional varieties of Portuguese spoken in the Upper Rio Negro. This term may be of Arawak origin; according to Koch-Grünberg (1906b: 877) the term was first used by Arawak peoples with reference to the original nomadic population, cf. Baniwa-Kurripako ma-aku (negation-speak) ‘those who cannot speak’ (also see Weir 1984: 15).

Makú languages considered here should not be confused with the following other groups to which this name is sometimes applied:

(i) Makú, or Maku, an endangered language spoken by a few old people on the river Auarí, a tributary of the Uraricoera which flows into the Rio Branco, Roraima, Brazil (see Migliazza 1985) — this is briefly discussed in §1.6 of chapter 13 below;
(ii) Makú, also known as Cofan-Makú, spoken around the lake of Cuyabeno, in Colombia and Ecuador (this falls outside Amazonia);
(iii) Makú, also known as Sáliba-Maco, or Maco-Piaroa, which is a group of the Sáliba-Piaroa family (see §3 of chapter 13 below), spoken in the savannas between the lower Ventuari and Orinoco in Venezuela.
They have traditional economic ties with Arawak- and Tucano-speaking peoples. Arawak- or Tucano-speaking subclans usually enter into a 'master-underling' relationship with a particular group of Makú (hence denominations of the type: Makú of Kubeo or Makú of Tucano). They lack the 'obligatory' multilingualism associated with 'linguistic' exogamy, so striking among the Tucano- and Tariana-speaking peoples of the Vaupés.

Makú-speaking men usually have some knowledge of the Tupinambá-based creole Língua Geral (in the region of Middle and Upper Rio Negro), of Tucano (in the Vaupés), or of Portuguese and/or Spanish. Women and children tend to be monolingual.

The Makú are believed to be the original population of the Upper Rio Negro (Nirnuendajú 1982: 169; Galvão 1979: 147), subsequently conquered by Arawak-, and then by Tucano-speaking peoples. According to the oral tradition of the Desana, Tuyuca, Kubeo, Tariana and Baniwa peoples, a few Makú groups actually lost their own languages and started using languages of their 'conquerors'; these 'former Makú' usually have a lower status in the internal hierarchy of their tribes (see Koch-Grünberg 1906b: 878; Janet Barnes p.e.).

**Table 9.1 The Makú languages in Brazil and Colombia**

<table>
<thead>
<tr>
<th><strong>Middle Rio Negro (Brazil)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Nadeb-Kuyawi (Nd-Ku)</td>
</tr>
<tr>
<td>Nadeb (Nd) is also known as Nadobó, Anodób, Makunadôbo, Guariba, Guaribatapuyó, Kabori, or Cabori, and Xiriwai. There are around 400 speakers mainly on the Unetiwi river - a tributary of the Middle Rio Negro - and along the Japurá river. Kuyawi (Ka) has around 20 speakers in the village Bom Jardim, on the south shore of the Middle Rio Negro.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>The Upper Rio Negro and Vaupés (Brazil and Colombia)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Dáw (D) (regional pejorative name Kamú)</td>
</tr>
<tr>
<td>There are around 83 speakers of Dáw, currently living in one community near São Gabriel da Cachoeira, Brazil.</td>
</tr>
<tr>
<td>3 Hupda-Yuhup (Hu-Yu)</td>
</tr>
<tr>
<td>Hupda (Hu) is known as Makú of Tucano and is spoken by around 1,000 people, occupying a considerable area between the rivers Papuri and Tiiquié in Brazil and Colombia. Yuhup (Yu) is spoken by around 400 speakers in 10 small communities between the Upper Rio Negro, the Tiiquié and their tributaries, and in the region of the Upper Japurá, Apaporis and Trairá, and Caquetá, in Brazil and Colombia.</td>
</tr>
<tr>
<td>4 Kakua-Nukak (Kk-Nk)</td>
</tr>
<tr>
<td>Kakua (Kk) is also known as Bará, or Makú of Kubeo, or Makú of Guanano. There are approximately 220 speakers, along the lower Vaupés and the Papuri, in Colombia near the Brazilian border. Nukak (Nk) is spoken by around 300–600 speakers scattered in jungle regions between the Guaviare and Iribí rivers, in Colombia.</td>
</tr>
</tbody>
</table>

Most of the Makú of the Middle and Upper Rio Negro were contacted by white people in the early twentieth century. Some groups, e.g. the Nukak, were completely unknown to the outside world until the 1980s.

**GENETIC CLASSIFICATION**

The Makú languages are listed in table 9.1. Abbreviations used in this chapter are given after each name.

The genetic unity of Makú languages was suggested by Koch-Grünberg (1906b), Rivet and Tastevin (1920), and then by Nirnuendajú (1950/5, 1982). Recent comparative studies show affinities in pronominal paradigms (but not in other grammatical markers); about 300 cognates have been established across the family (V. Martins forthcoming). Examples of lexical cognates are:

- Nd, Ku ʔib, Yu, Hu, Kk ʔip, D, Nk ʔiŋ 'father';
- Nd, Ku ʔib, Yu, Hu, Kk ʔip, D ʔiŋ ‘egg’;

Raised figures in Kakua examples indicate tones.
from Tucano into the Makú of the Upper Rio Negro may be the reason why these languages are so different from Nd and Ku.

There are enough differences between Nd-Ku and the others to establish them as a distinct group within the family. There are two possible family trees, given in diagrams 9.1 and 9.2; detailed comparative work will be needed to decide between these.


3 PHONOLOGY

Consonantal systems vary in their complexity, from 17 phonemes in D and 18 in Nd, to 12 in Hu-Yu and 12 in Kk. See table 9.3.

Nasalization is a syllable prosody in Hu-Yu and in Kk-Nk, so that blm, dln and glp are allophones. D and Hu-Yu have prenasalized stops as word-final allophones of simple stops.
Table 9.3 Consonants in Makú languages

<table>
<thead>
<tr>
<th></th>
<th>In all languages</th>
<th>D also has:</th>
<th>Hu-Yu also has:</th>
<th>Nd also has:</th>
<th>Kk also has:</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>p, b, t, d, k, g</td>
<td>t</td>
<td>f</td>
<td>f, t</td>
<td>t</td>
</tr>
<tr>
<td>fricative</td>
<td>h, f</td>
<td>h</td>
<td>f</td>
<td>f, t</td>
<td>t</td>
</tr>
<tr>
<td>affricate</td>
<td>tʃ, dʒ</td>
<td>tʃ</td>
<td>dʒ</td>
<td>dʒ</td>
<td>dʒ</td>
</tr>
<tr>
<td>nasal</td>
<td>m, n</td>
<td>m, n</td>
<td>n, n, p, g</td>
<td>n, n, p, g</td>
<td>n, n, p, g</td>
</tr>
<tr>
<td>liquid</td>
<td>w, y</td>
<td>w</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>glide</td>
<td>l</td>
<td>l</td>
<td>r [r], [d]</td>
<td>r</td>
<td>r</td>
</tr>
</tbody>
</table>

Table 9.4 Vowels in Makú languages

<table>
<thead>
<tr>
<th></th>
<th>In all languages (except Kk)</th>
<th>D also has:</th>
<th>Hu-Yu also has:</th>
<th>Nd also has:</th>
</tr>
</thead>
<tbody>
<tr>
<td>i, í, u, ü</td>
<td>i, í</td>
<td>i, í</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>e, o</td>
<td>e, o</td>
<td>e, o</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>ñ, ñ</td>
<td>ñ, ñ</td>
<td>ñ, ñ</td>
<td>ñ</td>
<td></td>
</tr>
<tr>
<td>a, à</td>
<td>a, à</td>
<td>a, à</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Table 9.5 Tones and stress in Makú languages

|                  | Nd-Ku: no tones, non-contrastive stress on the final syllable | D: two tone contours; almost predictable | Kk-Nk: four tones | Hu-Yu: not known |

Nd, D, and Hu-Yu have very complicated vowel systems, quite atypical for northwest Amazonia; see table 9.4. In contrast, Kk has a 'typical Amazonian' five-vowel system (i, e, a, o, plus high central i), just like many Tucano languages. Nasal vowels are phonemic only in D and Nd.

Syllable structure is CV(C); D and Hu-Yu have vowel assimilation at morpheme boundaries. Velar stops tend not to occur in word-initial position.

D has a low tone contour in 70 per cent of CVC syllables with a voiced coda; all the syllables with a voiceless coda have a high tone (the vast majority of words are monosyllabic). Tonal alternations may indicate a change in word class, e.g. weːd (low) 'eat', weː:d (high) 'food'.

Proto-Makú probably did not have any tone or pitch accent; these could have developed under the areal influence of Tucano languages, or as an independent innovation (V. Martins forthcoming).

4 WORD STRUCTURE

The Makú languages of the Upper Rio Negro and Colombia are agglutinating and predominantly suffixing with a few elements of fusion across morpheme boundaries. D and Yu have vestigial prefixes; in Kk just the subject (A/S) is cross-referenced on the verb with a prefix. In contrast, Nd has up to nine prefix positions.

Nd prefixes are thematic, aspect, valency-changing, locational and subordinating. Weir (1984: 50) reports that their ordering appears to be determined by their phonological form. Some of these – locational and subordinating, possibly some derivational and thematic – developed from incorporated postpositions and adverbials. There is also a set of cross-referencing proclitics. (1) shows a typical verb in Nd (Weir 1984: 51).

(1) ga-na-ni-wan
THEME-THEME-NEG + ASPECT-spill
'It does not spill.'

Another typologically unusual property of Nd is the position of incorporated adverbials, nominals and adpositions before the root (see (11)). Many other languages of the Amazon tend to postpose incorporated nominals/adverbials (see (n) in §2 of chapter 1 above).

Fossilized prefixes and traces of preverbal incorporated adverbs are found in D, e.g. kes-pij (inside-be.full) 'overcome'; faf-fin (inside-count) 'think', lit. 'count inside the head'; cf. Hu (Moore and Franklin 1980: 3): hì-y-kàd (underneath-pass) 'pass underneath'. A reflex of a prefixed noun marker in Nadeb is found in Dàw mar 'inherent possession', and in negative predicates mài and mëh.

This may indicate that proto-Makú had more prefixing tendencies than the modern languages of the Upper Rio Negro; it is quite likely that these languages acquired a predominantly suffixing profile under the areal pressure from East Tucano.

5 WORD CLASSES

The only open word classes are verbs and nouns. In D and Hu-Yu adjectival concepts are expressed with stative verbs, while in Nd 'adjectives' are a subclass of obligatorily possessed nouns (Weir 1984: 84). There is a closed class of positive and negative copula verbs used in equational, attributive and locative-existential clauses.

For instance, a prefix 7ih- is used to form names, e.g. wa-th 'pigeon', 7ih-wa-th 'pigeon' (proper name).
Alienable and inalienable possession is distinguished in Nd, and D and Hu-Yu, but it is marked in different ways.

Nd has three classes of nouns (Weir 1984: 83-7): (a) nouns which can only occur with a preposed possessor, e.g. Subih nnoh 'Subih's hand'; (b) those which can optionally occur with a possessor, e.g. Subih t'ab 'Subih's house'; (c) those which require a classifier when possessed, such as animals, plants, some kinship terms, etc., e.g. Subih waa maseel (Subih POSSV.CL:FOOD banana) 'Subih's banana' (one cannot say *Subih maseel).

In D and Hu-Yu both alienable and inalienable possession is marked by juxtaposition (Possessor-Possessed), e.g. D yiim dU/1/ (dog tai/) 'dog's tail'.

D is unusual in employing 'possessor classifiers' – suffixes which attach to possessors in alienable possession constructions; which is used depends on whether the possessor is animate or inanimate: -de?:'inanimate possessor', -ej 'animate possessor': see (2) and (3).

(2) yud daw-tog-ëj
   clothing human-female.child-CL:ANIM.POSSR
   'The clothing is a girl's', or 'girl's clothing'
(3) yak kaw-wa7-de?:
   manioc garden-up-CL:INAN.POSSR
   'manioc from a garden'

D and Hu-Yu also have generic noun classifiers in noun phrases, e.g. D daw tog (CL:HUMAN girl) 'a girl'. These are only rarely found in Amazonian languages.

D and Hu-Yu (but not Nd) have very productive nominal compounding; nominal compounds are sometimes hard to distinguish from classifier NPs, e.g. D daw-fob-ked (person-hand-inside)'the inside part of a hand'; daw-iim-bxk-jëen (person-eye-skin-hair)'eyelash'. These examples have just one stress, on the last syllable; this shows that they form one phonological word (unlike classifier-noun constructions).

Nd has an extremely complicated verbal structure with up to nine prefix positions. Prefixes are classified into six types (Weir 1994: 293-4):

(i) the formative prefix a-, attached to the verb root in the absence of other prefixes;
(ii) aspect prefix i- (obligatory with some verbs);
(iii) aspectual and modal prefixes;
(iv) thematic prefixes;
(v) valency-changing prefixes;
(vi) subordination prefixes.

D and Hu-Yu have a much less complicated verbal morphology. There is a great deal of verb compounding (root serialization) used for encoding aspectual meanings,
as in (6), from D, where the verb 'be straight' shows the immediate character of the action. (Up to five verbs can be compounded.)

(6) yo:y ba-ham-yow
    medicine spill-go-be.straight
'The medicine spilt straight away.'

D and Hu distinguish three tenses: present (unmarked), past (D -el) and future (D -ey).

In Nadeb every verbal root and suffix distinguishes indicative from non-indicative. Non-indicative forms are used in imperatives, nominalizations and interrogatives; they are marked with vowel change and/or the voicing of a final consonant (Weir 1984: 39). Some verbs also mark whether the S/A argument is singular or plural. Plural verb forms can be suppletive; or derived with a prefix or vowel and/or consonant alternations. There are only a few suffixes (diminutive, ingressive and completive; Weir 1984: 45).

While Hu and Kk have only one imperative (Hu -kém, Kk reduplication of the final vowel), Nd distinguishes several imperatives (simple, indirect, permissive). D distinguishes imperative (5h) and prohibitive (li).

**8 VALENCY-CHANGING DERIVATIONS**

D, Hu-Yu and Kk have no valency-reducing derivations. In contrast, Nd has a prefix ka- 'reflexive, reciprocal' (Weir 1984: 47). It is also used as an intransitivizer, and to mark non-agentive passive (1984: 107).

Nd also has a morphological causative, marked with da-. The only example given by Weir (1984: 47) is with an intransitive verb. Another technique for marking causatives is verb compounding as illustrated in (7) (1984: 44).

(7) tsé Subih na-boh-yat
    pig Subih THEME-shoot-lie down
'Subih killed a pig.' (lit. 'he shot; it is lying down')

In Kakua, morphological causatives (suffix -' - Cathcart 1979: 41) can be formed on intransitive - and probably also on transitive - verbs, while D and Hu use periphastic constructions, as in (8) and (9), both from D:

(8) tih ?ip we:d wa:y tih-tiy?
    he father eat order be-O.TOPICAL
    'His father made him eat.'

(9) ah do? haw na-jt-y?
    1sg causative boil water-O.TOPICAL
    'I made the water boil.'

D has a reported/non-reported evidentiality contrast. Evidentiality is lacking from the other languages.

**9 INCORPORATION**

A striking property of Nd (not found elsewhere in the family) is noun incorporation of two kinds. Similarly to other Amazonian languages (e.g. Guahibo) S/O can be incorporated only if obligatorily possessed, and the incorporated noun is placed before the verb. The original possessor becomes the object (Weir 1990: 323ff.); animate nouns rarely incorporate. The effect of incorporation is 'discourse' prominence of the new object. Compare (10) and (11):

(10) Subih tab iih ta-ma
    Subih house 1sg THEME-make
    'I am making Subih's house.'

(11) Subih iih tab-ta-ma
    Subih 1sg house-THEME-make
    'I am making a house for Subih.' (lit. 'I am house-making Subih'; the benefit for Subih is emphasized)

Nd allows more than one noun to be incorporated (similar to Cuiba, from the Guahibo family - see chapter 13 - but unlike the majority of Amazonian languages). Compare (12), lacking incorporation, with the incorporation in (13). It is also possible to incorporate two or even three nouns as in (14) and (15), but this is not common. The incorporated nouns are underlined. These examples are from Weir (1990).

(12) a hoonh tab nooh kad ga-jaa
    2sg+POSSR grandmother house mouth uncle THEME-close
    be.suspended
    'Uncle closed the door of your grandmother's house.'
(13) a hoon tob kad nooh-ga-joo
2sg + POSSR grandmother house uncle mouth-THEME-close
dak be.suspended
'Uncle closed the door of your grandmother's house.' (lit. 'Uncle mouth-closed your grandmother's house'; the effect on the door is emphasized)

(14) a hoon kad tob-nooh-ga-joo
2sg + POSSR grandmother uncle house-mouth-THEME-close
dak be.suspended
'Uncle closed the door of your grandmother's house.' (lit. 'Uncle house-mouth-closed your grandmother'; the effect on the house is emphasized)

(15) om kad hooñ-tob-nooh-ga-joo
dak 2sg uncle grandmother-house-mouth-THEME-close be.suspended
'Uncle closed the door of your grandmother's house.' (lit. 'Uncle grandmother-house-mouth-closed you'; the effect on grandmother is emphasized)

Adverbs and adpositions can also be incorporated. Unlike some other Amazonian languages, these are preposed to the verbal root. Incorporation of adpositions results in applicative-like derivations: if the verb is intransitive, the argument of the postposition becomes O, and the original S becomes A. If the verb is transitive, the argument of the postposition is promoted to O, and the old O is demoted to the periphery (Weir 1990: 326ff.). Consider (16) and (17). The incorporated postposition in (17) is underlined.

(16) eeA a-hing hxoół go
father FORMATIVE-go downriver canoe in
'Father goes downriver in a canoe.'

(17) hxoół eeA ga-hing
canoe father in-go downriver
'Father goes downriver in a canoe.' (lit. 'Father goes downriver-in a canoe')

This affects the discourse properties of 'canoe'. Incorporation of a postposition can be a strategy for topicalization (Weir 1984: 136); the transitivity of the verb is then not affected.

9 Makù

Table 9.6 Ergative cross-referencing in Nd

<table>
<thead>
<tr>
<th>S/O</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>2sg òm</td>
<td>ma- ~ òm</td>
</tr>
<tr>
<td>3sg ò</td>
<td>ta-</td>
</tr>
<tr>
<td>2pl</td>
<td>booh</td>
</tr>
<tr>
<td>3pl ò</td>
<td>la-</td>
</tr>
</tbody>
</table>

Table 9.7 Cross-referencing proclivities in Nd

<table>
<thead>
<tr>
<th>S/O</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>2sg òm</td>
<td>ma- ~ òm</td>
</tr>
<tr>
<td>3sg ò</td>
<td>ta-</td>
</tr>
<tr>
<td>2pl</td>
<td>booh</td>
</tr>
<tr>
<td>3pl ò</td>
<td>la-</td>
</tr>
</tbody>
</table>

(18) ta-kolâay ta-ta ta-mi-soo ta-kolâay me
3-claw 3-food 3-WITH+ASPECT-take 3-claw with
'As for his claws, it is with his claws that he takes his food.'

10 GRAMMATICAL RELATIONS

Nd is one of the most ergative languages in the Amazon region, while D and Hu-Yu are consistently nominative-accusative.

In Nd, an overt reference to A outside the predicate is obligatory, but the S/O argument may be omitted. This language also uses constituent order SV or VS and OAV or AVO, and it has a syntactic pivot S/O in coordination (Weir 1984: 89-91). There is also a split-ergative pattern for personal pronoun proclivities dependent on person and number — see tables 9.6 and 9.7.

D and Hu-Yu employ cases for marking core grammatical relations.9 D has an unusual pattern of core cases dependent on topicality of O and of A/S. This pattern may be the result of areal diffusion from Tucano languages. O is unmarked when it precedes the predicate as shown in (19).

(19) lakâhA mélO fōoh-ëp
hen jaca mim peck-PAST
'A hen pecked a jaca mim bird.'

8 Kk seems to use distinct cross-referencing markers for A, S and O; this needs further checking.

9 Kk seems to have a suffix (or enclitic?) which marks both O and addressee, -dhâ (Cathcart and Levinsohn 1976: 26-8).
When it is topical and follows the predicate, O is marked with -i:y7, as in (20).

(20) laka:h ʃɔː-h-eʔ ʃuː-t-iːy7
    hen peck-PAST jacamim-O.TOPICAL
    'As for the jacamim bird, a hen pecked it.'

The O suffix can be omitted when pragmatic context is likely to resolve any possible ambiguity. In a clause like (21), it is highly unlikely that a deer could kill a jaguar, so the object marker is usually omitted.

(21) ʔyánfįʔ ʃuː ʃoː-t-iːy7
    jaguar kill deer-O.TOPICAL
    'A jaguar killed the deer.'

A/S can be marked with an enclitic -tE:h if it is in contrastive focus and is highly agentive. In (22), the focus is on II, a mythological character.

(22) ʔaaten ʔil-teh wam ʃuː
    then II-AGENTIVE.A/S smoke much jacamim
    'Then it was II who smoked a lot of jacamim birds.'

Oblique constituents are marked with clitic postpositions. D is unusual in having an instrumental and three comitatives: -red 'instrumental', -did 'with' (equal relations between participants); -ray 'with' (one participant is superior to the others, e.g. chief, elder brother, etc.); -faʃ̃ 'with, among' (with inanimates: a mixture, e.g. coffee with milk). D, Hu-Yu and Kk have just one locative marker (D -hid, Kk -p 'to, from, in'), like Tucano languages. In contrast, Nd marks only ablative (bĮ) and locative (běit).

It is possible that proto-Makú could have been ergative, and that the languages of the Upper Rio Negro lost the ergative marking under the areal influence of Tucano languages. The main argument for this is the loss of ergative-type marking in constituent order in the languages of the Upper Rio Negro, accompanied by the loss of proclitics and prefixes.

Negation

All the Makú languages have complicated mechanisms for marking negation and have a special prohibitive marker. Nd has three negative morphemes: dooh, used to negate non-imperative main clauses (and as a negative answer); na, used to negate dependent clauses; and mani:h, used to negate imperative clauses (Weir 1984: 148-263; 1994). Besides simple negation, D has emphatic negation (contra-expectative) -taʔ, existential and possessive negation -tEh, and prohibitive -tEh.

Syntax

D has a number of coordinating and subordinating verbal enclitics, while Nd marks subordination and relativization with verbal prefixes. Hu seems to have a relativizing suffix (-pep) which goes on verbs (Moore and Franklin 1980: 15).

In Nd, only core arguments can be relativized. To relativize other arguments, they must be put into O position (via incorporation of a postposition). The relativizing morpheme appears within the relative clause, and the common argument is omitted from the main clause.

(23) [kapeh ʃuː iːh ʃuː jee doo] Subih
    coffee yesterday 1sg ASPECT-buy + INDIC REL Subih
    i-eek
    ASPECT-drink + INDIC
    'Subih is drinking the coffee which I bought yesterday.'

No such restrictions exist in other Makú languages – any argument in the main clause can be relativized.

Complement clauses in Nd are marked with subordinating verbal prefixes, while D uses simple juxtaposition of clauses. Only Kk has switch-reference, which has possibly developed through diffusion from Tucano languages.

Lexicon

D and other Makú languages of the Upper Rio Negro are highly 'verbal'; about 75 per cent of the words in texts are verbs. D has numerous verbs for 'carrying', 'cutting', 'falling' (depending on whether it is a person, or a fruit) and 'disappearing' (whether it is in the jungle or elsewhere). Similarly to East Tucano languages – but unlike Nd – D and Hu-Yu have just one word for 'sun' and 'moon'.

D has a fascinating system of counting. There is a word ʔmɛʔ 'one, alone', and tė:m 'pair' (used for two), and muː-wap 'three' (also used for 'many'). To say 'four', speakers show four fingers in twos saying ʔmɛʔ maab 'one brother' (each of the fingers has a 'brother'); 'five' is ʔmɛʔ maab meh 'one does not have a brother' (this goes on up to 10).
BIBLIOGRAPHY


Catheart, M. 1979. ‘Fonologia del cacua’, Sistemas fonológicos de idiomas colombianos, 266.


forthcoming. ‘Estudo histórico-comparativo da família lingüística Makú’.


The Nambiquara family consists of three languages:

1. Southern Nambiquara dialect complex
2. Mamainde/Nakarothe
3. Sabanes

The Southern Nambiquara dialect complex comprises of the following twelve intelligible dialects (superscript figures indicate tones – see below):

(a) Ne'su²
(b) Ki'tu²/hn²
(c) Ha'lot'te'su²
(d) Wa'ka/te'su²
(e) Wa'/sIh'su²
(f) A'kärte'su²
(g) Wa'ki/te'su²
(h) Ha'hai²/te'su²
(i) A'/yi/n?ka'kí/te'su²
(j) Ka'/hu'whí/ko'le'su²
(k) Tu'ka'/hu/la'kí/te'su²
(l) Wai'su²

The origin of the name Nambiquara is interesting. Lévi-Strauss (1948: 361) states:

"The Nambikwara (Nambikuara, Mambyuara, Malribarez) have been identified, only recently. Nambicuara, meaning 'long eared', was originally a Tupi nickname used since the 18th century for the little known tribes of the western and northern parts of the Serra dos Parecis. These tribes had large ear and lip plugs, like those of the Suya and Botaudo, and were called Beiros de Pau, 'Wooden Mouths', by the rubber collectors and gold miners. About 1830, they began to make hostile sorties from the region of the upper Sangue River. When, in 1907, General Candido Mariano da Silva Rondón discovered important tribes in the Serra do Norte, he identified them with the Nambikwara of the old literature. Thus, Nambicuara designates a tribe other than the 'Long Ears', or 'Wooden Mouths', to whom it was originally applied."

1 The author is grateful to his SIL colleagues Menno and Barbara Kroeker and Peter Kingston for the map of the Nambiquara area and for information on Mamainde. And also to the residents of the villages of Serra Azul, Camararé and Campos Novos for their patience in teaching him the language.

2 The name Nambicuara (alternatively spelt Nambicuara or Nambikwara) is interesting. Lévi-Strauss (1948: 361) states:

"The Nambicuara (Nambikuara, Mambiyuara, Malribarez) have been identified, only recently. Nambicuara, meaning 'long eared', was originally a Tupi nickname used since the 18th century for the little known tribes of the western and northern parts of the Serra dos Parecis. These tribes had large ear and lip plugs, like those of the Suya and Botaudo, and were called Beiros de Pau, 'Wooden Mouths', by the rubber collectors and gold miners. About 1830, they began to make hostile sorties from the region of the upper Sangue River. When, in 1907, General Candido Mariano da Silva Rondón discovered important tribes in the Serra do Norte, he identified them with the Nambikwara of the old literature. Thus, Nambicuara designates a tribe other than the 'Long Ears', or 'Wooden Mouths', to whom it was originally applied."
The Northern Nambiquara subgroup consists of two languages: the mutually intelligible dialects (m) Mamainde and (n) Nakarothe, and the single-dialect language (o) Sabanes.

Map 9 shows the approximate locations of the dialects of Nambiquara. These are numbered 1 to 12 and are given below. After each location name is the approximate number of residents there, followed by the dialects spoken there according to the labelling (a) to (o) used above. All the locations are within the State of Mato Grosso do Norte, Brazil, and are bounded by the Rio Papagaio to the east and Rio Guaporé to the west, both of which flow northward and are eventual tributaries of the Amazon. (The tributary sequences are Papagaio-Juruena-Tapajós-Amazon and Guaporé-Madeira-Amazon.)

1 Campos Novos, 48 (a, b)
2 Camararé, 73 (b)
3 Posto Nambiquara, 182 (c)
4 Utiariti, 62 (c, d)
5 Galera, 64 (e)
6 Fazenda Estrela, 82 (f, g)
7 Fazenda Zilo, 67 (h)
8 Quatro Pontos, 44 (e, h)
9 Sararé, 51 (i, j, k)
10 Mamainde, 103 (m)
11 Nakarothe, 25 (n)
12 Sabanes, unknown (probably less than 50) (o)

The description in this chapter is of Ki’tau’ihu, a dialect of Southern Nambiquara, and is based on some 2,000 pages of oral text and some 8 years of residence in Nambiquara villages. The dialect of Northern Nambiquara that has been extensively studied is Mamainde (e.g. Eberhard 1995). Its lower-level phonology and morphosyntax is not unlike that of the southern dialect which is to be the main concern of this chapter. Brief notes will be made from time to time in the description below when there are significant differences between the structures of the two dialects.

The first large-scale contacts of the Nambiquara with the outside world were with the Rondônia expedition of 1911. Estimates of their population at that time run from 20,000 to 50,000. Sadly, epidemics immediately after the Second World War reduced their numbers to less than 500 by the early 1950s. Happily, their numbers have increased significantly in the last 15 years; in some villages the population has more than doubled. Most of the population is young. Until about 20 years ago, the Nambiquara lived a traditional hunting and gathering life, supplemented by some slash-and-burn agriculture. There was little contact with the surrounding Brazilian culture, and only a few of the men spoke even rudimentary caboclo Portuguese.

Today many of the younger people speak good Portuguese. The children, however, still learn the vernacular and all Nambiquara use it.

### PHONOLOGY

The phonemes of Nambiquara are given in tables 10.1 and 10.2.

The (f) and (m) within parentheses can occur only in loan words. However, the nasal phone [n] and its prestopped counterpart [bm] can occur as variants of the phoneme /nl/. (The phonemic system of Mamainde has no contrastive implosive /dl/, and /ml/ is quite rare.)

There are three contrastive contour tones marked as 1 (down-gliding), 2 (up-gliding) and 3 (level). (Mamainde has a fourth tone which is at a lower level and which occurs quite rarely.) Nasalization (shown by ‘ over a vowel) is contrastive on vowels. Laryngealization (shown by _ under a vowel) is also contrastive (not merely conditioned by the presence of a glottal stop), on both oral and nasal vowels. Aspiration can occur contrastively on the voiceless stops /pl, /tl, /kl, on the nasal /nl, and the liquid /l/. Clearly the aspirated versions of these sounds could have been interpreted as five extra phonemes. But, in the interests of economy, they have been
interpreted as clusters with /hl/. This interpretation also leads both to more symmetrical distributional statements, and to simpler morphophonemics.

2.1 The variants of phonemes
The stops /pl, /tl, /kl/ have variants that freely fluctuate from voiced to voiceless in all environments: [p]—[b], [t]—[d], [k]—[g]. The nasal /nl, when occurring in the coda of a non-word-final syllable, has prestopped variants [bm]—[dn]—[gn] when the syllable nucleus is an oral vowel, and the whole stop-nasal sequence assimilates to the place of articulation of the initial consonant of the following syllable. When the nucleus is a nasal vowel, however, the variant is simply a nasal consonant [m]—[n]—[n] that again assimilates to the place of articulation of the initial consonant of the following syllable. The liquid /l/ has variants [l] and [r], the lateral preceding the back vowels /a, /o/ and the flap preceding the front and central vowels /i, /e/ and /u/. As for the vowels, the high back /u/ has a variant range over [u]-[o]-[i]-[a] in unstressed syllables. The low central /a/ is realized as [a] in unstressed syllables.

2.2 Syllable structure
The syllable structure is (C(CCC)V(C). By far the most common syllable type is CV. The optional final C slot can be filled only by the voiceless stops /l, /kl/, /fl/, the nasal /nl/ and the semivowel /nl/. As for the onsets, the first slot can be filled by any consonant in the inventory, but subsequent onset slots can be filled only by the glides /l/, /l/, /l/. The initial syllable of the following word is an example of a syllable with a maximum onset:

(1) [kwh?a?k?ia?liis?u] 'kind of deer'

2.3 Stress
Stress is realized as increased intensity, with an increased pitch range on the falling and rising glides 1 and 2, and with a lowered-level pitch on the level tone 3. It comes on the last syllable of a morpheme cluster.

2.4 Morphophonemics
The surface forms of both nouns and verbs follow a few morphophonemic rules.

For verbs there are various different surface forms of the 1sg and 2sg subject suffixes and the 3sg negative subject suffix. For nouns there are various different surface forms for the different nominal endings. The most economic way to describe these is by setting up underlying forms for the verb or noun roots (or stems) from which the various surface forms can be easily derived.

Specifically, we take the verb roots (or stems) to have underlying forms which can end in a vowel or in one of the consonants /l, /h, /l/, and the underlying forms of the 1sg and 2sg subject suffixes to be -a/- and -in/- respectively (as shown in table 10.7). Then, for example, the verb root with underlying form suP- 'hit' (ending in /l/) would derive the surface forms of its suffixes typically as follows:

<table>
<thead>
<tr>
<th>UNDERLYING FORM</th>
<th>SURFACE FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>suP-a-l-a2</td>
<td>hit-1sg-PERFV</td>
</tr>
</tbody>
</table>

where the crucial step in the derivation is resyllabification whereby the final closing consonant of the underlying form of the verb root becomes the initial consonant in the surface form of the 1sg (or 2sg) subject suffix. Hence the 1sg subject suffix will have surface forms -a/-, -bu/-, -ba/-, -ba/-, -ba/- depending on which consonant, if any, closes the final syllable of the underlying form of the verb root.

Similarly noun roots have underlying forms which can end in a vowel or in one of the consonants /l, /h, /l, /s, /k/ (see §3.3). The definite noun suffix in its underlying form is -a/-, and again we get the correct surface forms of the definite nouns by a derivation where the crucial step is a resyllabification. Here are some examples of definite nouns in their underlying and surface forms.

<table>
<thead>
<tr>
<th>UNDERLYING FORM</th>
<th>SURFACE FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>s7?lh2-a2</td>
<td>s7?l-ha2</td>
</tr>
<tr>
<td>house-DEF</td>
<td>'the house'</td>
</tr>
<tr>
<td>wát2-a2</td>
<td>wát2-la2</td>
</tr>
<tr>
<td>cloth-DEF</td>
<td>'the cloth'</td>
</tr>
<tr>
<td>kat1-a2</td>
<td>kat1-ta2</td>
</tr>
<tr>
<td>stick-DEF</td>
<td>'the stick'</td>
</tr>
</tbody>
</table>

Note that when the underlying form of the noun root ends in a voiceless plosive, that voiceless plosive is doubled in the surface form.

3 MORPHOLOGY
The morphology is largely agglutinative, but there are also some complex morpheme fusions in some of the verb forms. The language is head-marking and predominantly suffixing, but there are a few prefixes.
There are three open word classes in Nambiquara, verbs (§3.1), adjectives (§3.2) and nouns (§3.3), and four closed word classes, pronouns (§3.4), adverbs (§3.5), interjections (§3.6) and ideophones (§3.7).

### 3.1 Verbs

3.1.1 Main verbs

Main verbs are suffixed for mood, person–number and aspect. Mood can be either indicative or imperative, and indicative verbs are further suffixed for tense–evidentiality.

(i) Indicative verbs

Indicative verbs can be either declarative (asserting information without any degree of doubt) or dubitative (giving information but qualified by varying degrees of doubt).

Indicative main verbs are inflected for the following obligatory categories:

- subject person: first, second, third
- subject number: singular, dual, plural
- speaker number: singular, plural
- aspect: perfective, imperfective
- tense: future, present, recent past, mid past, remote past

(Mamainde has two futures, one of which is more definite than the other. It also distinguishes recent past, mid past and distant past.)

The non-future Nambiquara forms are further obligatorily inflected for:

- evidentiality: observational, inferential, quotative, internal support
- newness: given, new (with the usual meanings)

Evidentiality has to do with the kind of evidence that the speaker adduces to support their statement. The four main subcategories mentioned above have the following meanings:

- with observational support the speaker claims to have seen the action they are reporting;
- with inferential support, the speaker’s claim can be based either on seeing an associated simultaneous action and making an interpretation therefrom, or on seeing a set of circumstances which must have resulted from a previous action and making an inference; different suffixes mark these two options;

Table 10.3 Nambiquara new-information verbal suffixes

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>-wa²</td>
<td>imperfective aspect</td>
</tr>
<tr>
<td>-ra²</td>
<td>perfective aspect</td>
</tr>
<tr>
<td>-nii² ~ nia²</td>
<td>inferential</td>
</tr>
<tr>
<td>-ta²</td>
<td>observed circumstances</td>
</tr>
<tr>
<td>-na²</td>
<td>observed action</td>
</tr>
<tr>
<td>-ta³</td>
<td>quotative</td>
</tr>
<tr>
<td>-he³</td>
<td>mid past tense external</td>
</tr>
<tr>
<td>-hia²</td>
<td>present tense internal</td>
</tr>
<tr>
<td>-he²</td>
<td>recent past tense internal</td>
</tr>
<tr>
<td>-he³</td>
<td>mid past tense internal</td>
</tr>
<tr>
<td>-na³</td>
<td>action currently observed by both speaker and hearer</td>
</tr>
</tbody>
</table>

With quotative support the speaker is simply passing on information they have heard from another speaker;

With internal support the speaker reports their ‘gut feeling’ that that which they assert must be so.

(Mamainde has the distinctions of observational versus inferential versus quotative support, but there does not seem to be any distinction between given and new information expressed in the verb morphology.)

Table 10.3 gives the surface forms of some new-information forms with some forms showing typical category combinations.

Here are some typical examples of new-information verb forms:

(2) wa'kon²-na¹-ra²
work-ACTION.CURRENTLY.OBSERVED.BY.BOTH.S.&.H.-PERFV
‘He is working.’

(3) wa'kon²-na¹-ra²
work-OBSERVED.ACTION-PERFV
‘He worked.’ (I observed, recently)

(4) wa'kon²-tna²-he²-ra²
work-QUOTATIVE-MID.PAST.INTERNAL-PERFV
‘He worked.’ (I was told, in the past)

(5) wa'kon²-nin²-ta²-wa²
work-INFERRENTIAL-OBSERVED.CIRCUMSTANCES-IMPERFV
‘He worked.’ (I inferred from observed circumstances recently)
work-IMPERFV

"He is working." (I inferred from observed associated action from a distance recently)

Given-information forms have many fused morphemes, so I merely give a few representative forms here, without attempting any morpheme segmentation.

(7) wa3koJ-tai'ti'tu'wa^2

"He worked." (mid past, we both saw him do it)

(8) wa3koJ-te'ni'tu'wa^2

"He worked." (recent past, we both saw him do it)

(9) wa3koJ-te'nü'ti'tu'wa^2

"He worked." (recent past, we were both told)

(ii) Imperative verbs

Imperative main verbs are suffixed for:

- speaker number: singular, plural
- subject person-number: 2.singular, 2.plural
- aspect: perfective, imperfective

There are three different forms of positive imperative verbs, one expressing an action to be done in the immediate future, and two others expressing an action to be done in the more distant future. However, there is only one form of negative imperative verbs. (Mamainde has only two positive imperative forms, one strong, the other more gentle, almost a permissive. There is also one negative imperative form.)

In reported speech and reported thought quotes, the main verb takes on a different set of aspect markers from those used when the main verb is in an ordinary statement. Thus contrast:

(11) wa3koJ-na'tu'-wa^2

work-lsg-FUT-IMPERFV.STATEMENT

"I will work."'

(12) wa3koJ-na'tu'-wi^1

work-lsg-FUT-IMPERFV.QUOTE

""I will work", he said."

Table 10.4  Namibiquara subordinate verb inflections

<table>
<thead>
<tr>
<th>Verb-like inflections:</th>
<th>-nöl'na^1, -të'na^1, -ka'yt^1, -te'ka'ka^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 temporal sequence</td>
<td></td>
</tr>
<tr>
<td>2 conditional</td>
<td>-k?ar'na'nty^1 (highly probable),</td>
</tr>
<tr>
<td></td>
<td>-k?e'la'ce'k?at^1 (less probable)</td>
</tr>
<tr>
<td>3 adversative</td>
<td>-la^1</td>
</tr>
<tr>
<td>4 high-level switch</td>
<td>-ha'ha'li^1</td>
</tr>
<tr>
<td>Nominal-like inflections:</td>
<td></td>
</tr>
<tr>
<td>5 spatio-temporal-circumstantial</td>
<td>-täu'^1</td>
</tr>
<tr>
<td>6 causal</td>
<td>-ha'k?al^1, -jut'ku^1, -jau'ku'^1, -k?e'su'^2</td>
</tr>
</tbody>
</table>

3.1.2  Subordinate verbs

Subordinate verbs form the nucleus of subordinate clauses. They can only take indicative mood, and they are suffixed for subject person and number, and speaker number, exactly like indicative main verbs. But the remainder of the subordinate suffix strings are different. Some subordinate verbs end with nominal-like inflections, but others do not. We thus divide the categories of subordinate inflection into two subclasses. These inflections are given in table 10.4.

In the first subclass, that of the verb-like inflections, the subordinate verb consists of a verb root followed by person-number suffixes and finally a subordinate inflection. There is no possibility of any further inflection following the subordinate inflections. Specifically, none of the tense-aspect-evidentiality inflections found on main verbs, nor any of the definiteness-demonstrative-evidentiality-causality suffixes found on nominalizations, are possible on any of the subordinate verb forms of this first subclass. Nor do the subordinate verbs of this first subclass fill syntactic slots that are normally filled by nouns.

In the second subclass, that of the nominal-like inflections, the forms are quite different. There is firstly the verb root, followed by the person-number suffixes, and then by the nominal-like subordinate inflection. Specifically, of the four causal suffixes, the last three ju'l'su'^1, ju'ls'he, k?e'su'^2 all end in -su'^1, the indefinite ending for nouns. The last morpheme -k?at'^1 on the first causal ending -ha'k?at'^1 is also frequently found on nouns – for example,
In the text from which this example was taken, the speaker had been carrying his small child in his arms on a journey but both had been caught in a violent rainstorm. The father left his child at a shelter in the care of others and carried on to a village to fulfil an errand. Now, however, with his errand accomplished and the rain having also abated, the father feels he can attend to his child. This is when he says (14) 'The child therefore I will go back, pick her up and return.' In the sequence of things that he had to do, the time had logically come for him to deal with the child.

It is difficult to distinguish between the meanings of the temporal sequence suffixes. If the spatio-temporal-circumstantial option is chosen, there are further options of evidentiality possible. The meanings of the four causals have been discussed at length in Lowe (1990). They relate to mental motivation (e.g. I saw the rain coming, I thought that I didn't want to get wet, so I went home) and situational motivation (e.g. I feel hungry, so I will go hunting), among other factors.

Here is one example to illustrate the usage of each of the verb-like subordinate clauses.

1. Temporal sequence. Both the subordinate verb and the main verb describe actual events in time sequence in the story.

   (15) t̄a2no l-na l-ka3t1i2 a3-na l-he3_ra
           approach-1sg-SEQUENCE shoot-1sg-MID.PAST.
           EXTERNAL-PERFV
   'Coming up to it [the pig], I shot [it].'

2. Conditional. The conditional clause describes a hypothetical event or state of affairs. The main clause which almost always follows the conditional clause, describes a situation contingent upon the truth of the conditional.

   (16) t̄a2lú1_ai2Ji2 j3hit1-ta1_ta1 a3nú1_s?a3
           tapir-this chase-1sg-ADVERSATIVE dive-and
           a3?yoa2ha2ti2ha2ti2sai1ka3l2?
           swim.across go-and
           a3un1sa1-he2-ra2
           escape-1sgO-REC.PAST.INTERNAL-PERFV
   'I chased this tapir, but it dived (into the river) and swam across, and escaped from me.'

3. Adversative. In the following example, the subordinate or adversative clause comes first and describes a situation. The main clause follows and describes a situation contrary to the expectation of the first.

   (17) a3li1-a1i2 ji2hit1-ta1_l  ānú1-sa2l1
           plane-CL:FLAT.SHEET.LIKE-DEF come-descend-COND
           t̄a2?yu2sa1-lihun1-nha2-ra2
           rise-1sgO-want-PRES.INTERNAL-IMPERFV
   'If the aeroplane lands I want to fly [in it].'

   Note that the aeroplane is classified as flat-sheet-like because of its wings. (See §3.3 for more information on the use of classifiers.)

   In the preceding textual context of this example, the speaker has been talking about the plans of his two hearers. Now he switches over to talking about himself and what he plans to do. This latter topic takes up much of the discourse that follows.

3.2 Adjectives

Adjectives are inflected for the categories of tense, aspect and evidentiality just like third-person verbs. In mood they are limited to the indicative.

   (19) wain1-na2-ra2
           straight-ACTION.CURRENTLY.OBSERVED.BY.BOTH S.& H-PERFV
   'It is straight.' (or 'correct')

   (20) wi2-na2-ra2
           good-ACTION.CURRENTLY.OBSERVED.BY.BOTH S.& H-PERFV
   'It is good.'
Adjectives could be regarded as a subclass of verbs with certain limitations on their pronominal and mood inflections. To convey the information that a first- (or second-) person referent is the bearer of an attribute, the kind of construction is used:

(21)  nün'kunJ-te2-sal_wa2
      strong-ATTRIBUAND-lsgC-IMPERFV
      'I am strong.'

In (21) the lsgC is a first singular copula form (see table 10.7).

3.3 Nouns

Noun roots, like verb roots, may end in a vowel, or in one of the consonants n, t, h, l, s, k. Noun stems often involve a classifier suffixed to the root. The classifier often describes the shape of the noun referent but there are a few more abstract classifiers as well. A list of classifiers is given in table 10.5.

The last four classifiers listed in table 10.5 can only occur in nominalizations. All of the others can occur in both noun stems and in nominalizations from verbs. (Mamainde has classifiers covering essentially the same semantic categories but some of the Mamainde forms are very different.) Below are some examples of classifiers on noun stems. In each case the classifier comes immediately after the noun root.

<table>
<thead>
<tr>
<th>Table 10.5 Nambiquara classifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>kútÁ</td>
</tr>
<tr>
<td>suÁ</td>
</tr>
<tr>
<td>kíÁ</td>
</tr>
<tr>
<td>énÁ</td>
</tr>
<tr>
<td>yauÁ</td>
</tr>
<tr>
<td>n?ánÁ</td>
</tr>
<tr>
<td>ka?loÁ</td>
</tr>
<tr>
<td>rúñÁ</td>
</tr>
<tr>
<td>thínÁ</td>
</tr>
<tr>
<td>koÁ</td>
</tr>
<tr>
<td>jahÁ</td>
</tr>
<tr>
<td>ka?luÁ</td>
</tr>
<tr>
<td>yutÁ</td>
</tr>
<tr>
<td>jauÁ</td>
</tr>
<tr>
<td>káÁ</td>
</tr>
<tr>
<td>k?eÁ</td>
</tr>
<tr>
<td>teÁ</td>
</tr>
</tbody>
</table>

Classifiers could be regarded as a subclass of verbs with certain limitations on their pronominal and mood inflections. To convey the information that a first- (or second-) person referent is the bearer of an attribute, the kind of construction is used:

(22)  wa3IinJ-suJ-S1l2 
      manioc-CL:BONE.LIKE-INDEF 
      'manioc root' (indefinite form)

(23)  hukÁ-én1-suÁ 
      shooter-CL:HOLE.LIKE-INDEF 
      'shotgun' (the gun barrel is the hole)

(24)  hukÁ-kiÁ-suÁ 
      shooter-CL:ROUND-INDEF 
      'bow' (the arc of the bow is round)

Classifiers can be used as deverbal nominalizers: for example,

(25)  s?iÁ|haÁ ?yauÁ-ain1-thiÁ-na2 
      house  live-3non.sg-CL:HOUSE.LIKE-DEF 
      'The house they live in.'

(26)  wán'taÁ  eÁ|eÁ-kíÁ-sain1-jauÁ|aiÁ-naÁ 
      word speak-to-they.to.me-CL:WORD-this.definite 
      'This word that they spoke to me.'

Classifiers also occur on adjectives used as modifiers in an NP: for example,

(27)  waÁ|laÁ wÁ|winÁ-ka?loÁ-aÁ 
      cloth blue-CL:FLAT.SHEET.LIKE-DEF 
      'the blue cloth'

They are not used with an adjective when it is employed predicatively: for example,

(28)  waÁ|laÁ wÁ|winÁ-naÁ-raÁ 
      cloth blue-ACTION.CURRENTLY.OBSERVED.BY.both.s.&.ii-PERFV 
      'The cloth is blue.'

Classifiers are also used with numerals: for example,

(29)  aÁ|laÁ soÁ|liÁ knÁ|haÁ|liÁ huÁ|tiÁ-taÁ-héÁ-raÁ 
      parrot only CL:ROUND-TWO SHOOT-1sg-MID.PAST.EXTERNAL-PERFV 
      'I shot only two parrots.' (parrots are viewed as round objects)

Classifiers also figure frequently in the syntax and will be referred to again in §4.5.

Nouns can be suffixed for definiteness and causality. Definiteness has three values - indefinite (-saÁ), or definite (-aÁ) or conditional (-aÁ). Indefinite nouns can be
Further inflected for causality but for no other categories. Definite nouns, however, can be further inflected for demonstrativeness, spatio-temporality, evidentiality and causality. However, only a limited number of combinations of time and evidentiality have been observed even in a very large corpus of nouns. For evidentiality, three categories, observational, inferential and quotative, have been observed on nouns, but for newness, only given-information inflections occur. Observational evidentiality seems to exclude 'current' (where current means 'at the time and place reached in the discourse'). And curiously, in folktales it is the inferential endings which are the ones most frequently found on nouns (where one might expect that a quotative evidentiality was more appropriate). The endings that have been found on Nambiquara definite nouns are given in Table 10.6.

Since the glosses on these suffix sequences may appear rather abstract, here are a few examples of their use on actual nouns:

(30) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{a}^\text{1}\text{a}\]
manioc-CL:BONE,LIKE-DEF
'the manioc root'

(31) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{ai}^\text{2}^\text{na}\]
manioc-CL:BONE,LIKE-DEF,CURRENT
'This manioc root which we both see before us now.'

(32) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{ni}^\text{3}\text{t}\text{i}\]
manioc-CL:BONE,LIKE-OBSERVATIONAL,RECENT,PAST,GIVEN
'The manioc root that both you and I saw recently.'

(33) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{ai}^\text{2}^\text{tal}\text{a}\]
manioc-CL:BONE,LIKE-OBSERVATIONAL,MID,PAST,GIVEN
'The manioc root that both you and I saw some time past.'

### Table 10.6 Nambiquara nominal endings

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>-a\text{1}</td>
<td>definite, unmarked</td>
</tr>
<tr>
<td>-ai\text{n}a\text{2}</td>
<td>definite, current</td>
</tr>
<tr>
<td>-in\text{ti}\text{3}</td>
<td>observational, recent past, given</td>
</tr>
<tr>
<td>-ai\text{tu}^\text{ti}\text{2}</td>
<td>observational, mid past, given</td>
</tr>
<tr>
<td>-ni\text{li}^\text{a}</td>
<td>inferential, definite, unmarked</td>
</tr>
<tr>
<td>-ni\text{li}^\text{a}^\text{na}^\text{2}</td>
<td>inferential, current</td>
</tr>
<tr>
<td>-au\text{t}^\text{e}^\text{ti}\text{2}</td>
<td>quotative, mid past, given</td>
</tr>
</tbody>
</table>

### Table 10.7 Nambiquara pronouns

<table>
<thead>
<tr>
<th>Subject</th>
<th>Object</th>
<th>Copula</th>
<th>Free</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>-1\text{1}</td>
<td>-sa\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>1non.sg.incl</td>
<td>-ki\text{1}</td>
<td>-ni\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>1non.sg.exc1</td>
<td>-ni\text{1}</td>
<td>-sa\text{a}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>2sg</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>2pl</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>3sg.masc</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>3sg.fem</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
</tr>
<tr>
<td>3non.sg</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
<td>-1\text{1}</td>
</tr>
</tbody>
</table>

(34) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{ai}^\text{2}^\text{t}^\text{2}\]
manioc-CL:BONE,LIKE-OBSERVATIONAL,MID,PAST,NEW
'The manioc root that I saw (but you didn't) some time past at some distant place.'

(35) \[\text{wa}^\text{lin}^\text{su}^\text{a}^\text{1}-\text{ni}^\text{3}\text{t}^\text{2}\]
manioc-CL:BONE,LIKE-INFERENTIAL,DEF,UNMARKED
'The manioc root that must have been at some time past, as inferred by me (but not by you).'

### 3.4 Pronouns

The suffix pronouns and the free pronouns are given in Table 10.7. Note that second person distinguishes singular, dual and plural, while first and third persons have just two numbers, singular and non-singular (covering dual and plural).

The reflexive pronoun suffix is -\text{n}a\text{2} and the reciprocal is -\text{n}ya\text{2}, both valid for all persons.

### 3.5 Adverbs

There is really no well-defined class of adverb words. However, adverbial functions can be realized by adjectival roots prefixed to a verb stem. Such adjectival roots are true prefixes; they obey word-internal morphophonemic rules. For example:
Some examples are hai ‘Right!, OK!’, ta’wi ‘No way!’ As expected, these forms do not take affixes, and do not conform to the regular phonological template of the language.

3.7 Ideophones

Ideophones are words whose sounds depict the action they refer to. Examples are [ka’lagn3] ‘sound of a heavy fall’ and [taubm l] ‘sound of footsteps’.

4 Syntax

The unmarked constituent order for both main and subordinate clauses is AOV, SV. The grammatical relations A, S, O are marked by cross-referencing pronominal suffixes on the verb: A and S being marked by subject pronouns, O by object pronouns. Topicality considerations change the unmarked word order so that the topic comes initially.

4.1 Main clauses

Main clauses can be intransitive motion, intransitive non-motion, stative intransitive, descriptive, existential, transitive.

Intransitive motion clauses describe the motion of the subject as in:

(38) ka’i’na’h2ta’n2a1 á’n2i’-h2a1-ra2
this.morning run-1sg-PERFV
‘This morning, I ran.’

Intransitive non-motion clauses describe the intransitive non-motion action of an agentive subject as in:

(43) daPna 2 a2hoh 3JP-sal -wa2
I sg.DEF.CURR expert-lsg-copula-IMPERFV
‘I am an expert.’ (a2hoh JliJsu2 is the underlying noun ‘an expert’)

Transitive clauses have the unmarked constituent order AOV as, for example, in:

(44) daPni2 d’luh 1 -ai2na2 té3·al·tu l-wa2
1sg.VAGUE.CURR woman-DEF.CURR take-I sg-FUT-IMPERFV
‘I will take (marry) this woman.’

However, discourse considerations often front the object NP and give an OAV order. There are, in fact, no case markings of any sort on the core syntactic NPs to distinguish functions such as A, S and O. However, in a few instances where the subject also has high topicality, the ‘highly topical subject’ suffix, -h2?Pserrsu!, is used, as in:

(45) we3-sa2 kwª3lha 2_kªunJ-jah 1·lait 1ta2
child-DEF name-young-male-MID.PAST-HIGHLY.TOPICAL.SUBJECT

Stative intransitive clauses describe the state experienced by a patient subject as in:

(40) dai?ri2 kat’jan’jah’lai’na2 yuh’lii’-sa1-nha2-wa2
1sg.VAGUE.CURR white.man fear-1sg-O-PRES.INTERNAL-IMPERFV
‘I am afraid of the white man.’

Descriptive clauses describe an attribute of an individual (or attribuand) as in:

(41) sa?we’na2 wi l-na’ra3
forest good-is
‘The forest is good.’

Existential clauses assert the existence of something as in:

(42) g’leñ’su2 na’h2人数2
armadillo.hole was
‘There was an armadillo hole.’

Copula clauses do not have an overt surface copula verb as such, but the copula relationship is expressed by suffixing the noun root predicate with the copula pronouns (see the third column in table 10.7). For example:

(43) daPni2 a’hoh’lii’-sa1-wa2
1sg.DEF.CURR expert-lsg-copula-IMPERFV
‘I am an expert.’ (a’hoh’lii’su2 is the underlying noun ‘an expert’)

(44) dai?ri2 d’luh 1 -ai2na2 té3·al·tu l-wa2
1sg.VAGUE.CURR woman-DEF.CURR take-I sg-FUT-IMPERFV
‘I will take (marry) this woman.’
Free translation: 'It was Mr Kwãplha2 the younger who murdered my child.'
Literal translation: 'The child, it was Mr Kwãplha2 the younger who murdered him of me.'

(Note incidentally that the present tense and the 'observed by both speaker and hearer' evidentiality is used by the speaker to report an action that neither the speaker nor the hearer ever saw! This is a rhetorical device to make the report more vivid.)

There are also many other cases in which the preferred strategy is not to use a single clause with two third person NP constituents, but rather two successive clauses with a repeated verb, introducing first the A of the action and then the O in the two clauses. For example,

(46) ä2nü2-a2  ka3la2-ti21  äni2-so21-ti2-hyu1-nü2la2  people-DEF many by-hand-gather-they.to.him-SEQUENCE 'Many people gathered up (the corpse) by hand,'
?yo2wen2-kah1-ha2-so21-hyu1-nü2la2  name-male-DEF-son-DEF murder.they.to.him-O -nü2la2-k?ai2lu2  -INFERENTIAL.DEF.UNMARKED-therefore so21-hyu1-nü2la2  gather.they.to.him-SEQUENCE they gathered up Mr ?yo2wen2's son who had been murdered, and sa2nü2-k?ii2-nhyain1-ta1-hê2-ra2  bury.to-they.to.him-QUOTATIVE-MID.PAST.INTERNAL-PERFV they buried him (it was reported).'

Ditransitive clauses: there are certainly verbs like 'give' and 'speak' that can take three non-oblique arguments, but almost never in text do we find a single clause with three overt NP constituents. Rather, we find the information spread over two or more clauses as, for example, in:

(47) ?yo2wen2-kah1-ha2-so21-k?ai2lu2  name-male-son-DEF.therefore Mr ?yo2wen2's son

In this example, the simplest translation into English would be in the form of a single ditransitive clause 'Mr ?yo2wen2 says to Mr kwãplha2 "content of quote".' There is the verb 'say', the subject (Mr ?yo2wen2), the indirect object (Mr kwãplha2), and the direct object ('these words' plus the content of the quote). However, in Nambiquara, the same information is invariably spread over two clauses as above. The first clause introduces the speaker and addressee, and has the verb 'say' but it is ambiguous as to who the speaker is. The second clause has a verb phrase 'he take and he say to him', and the subject of this verb phrase is Mr Y, and the object is 'words'. It is now finally clear who the speaker is.

4.2 Subordinate clauses

Subordinate clauses have the same constituent order as main clauses, and the verb inflections are as already set out under subordinate verbs. Relative clauses consist of a clause whose nucleus is a nominalization. Either the subject or the object of the embedded clause can be relativized. There is no surface relative pronoun form as such. The relative clause usually modifies a noun head, as shown in the following examples:

(48) ä2nü2-a2-nü2tal2  oh'nau2a2  person-DEF-INFERENTIAL.DEF.UNMARKED above y?au1-k?ii2-jah1-lo2-nü2tal2  stay.to-male-INFERENTIAL.DEF.UNMARKED 'the person who stayed up high' (subject relativization)
4.3 Clause coordination

Clause coordination is effected by using either one of the coordinate endings, -í' or -ha 2 k7aPi, in place of the tense-aspect endings on the verbs in each of the coordinate clauses. Then, in main clauses, the tense-aspect-evidentiality for the whole coordinate construction is expressed by one or more suffix strings at the very end of the coordination. Thus, for example:

(52) wa2la2 wa2hiJ I-F wa2hoJ7-i2-nal-tfíl-í'al
   clothes wash-coORDINATE bathe-coORDINATE-lsg-FUT-IMPERFV.THOUGHT
   na l-he 2-ra2
   lsg-REC.INTERNAL-PERFV
'i intend to wash my clothes and to take a bath.'

(Here the suffix strings na l til7a /¡€2 r {¡2 express the tense-aspect-mode-­evidentia­lity for the whole coordination.)

(53) nUn'ntiSJ2 su2-lal-1-1ha'k?ai3 su'naJ2-niSJ2
   lizard-defREM kill-1sg-COORDINATE armadillo-defREM
   su2-lal-1ha'k?ai3 na l-he 2-ra2
   kill-1sg-COORDINATE lsg-MID.PAST.INTERNAL-PERFV
   'I killed a lizard and an armadillo.'

It is difficult to get any lower-level distinction in the meanings of these two coordination endings. Their discourse functions are rather different in that the -ha'k?ai
4.5 Nominalizations

All classifiers may, as mentioned in §3.3, also function as nominalizers:

(60) s?1-ain1-thi3-na2
    stay-3n01.sg-CL;HOUSE.LIKE-DEF
    'the house (or village) where they stay'

As mentioned before, there are four classifiers which do not serve to derive noun stems from noun roots, but which only function as nominalizers, for example,

(61) 3?yen1-k?1a2 e1k?1-san1-jut3-ta2 3-sha1-tu1-wa2
    things-DEF speak-you.to.me-SITUATION-DEF see-1sg-PUT-IMPERFv
    'The situation of the things you spoke to me about, I'll take a look at.'

Classifiers figure quite prominently in the syntax. They can function like 'anaphoric pronouns' to refer back to an item mentioned previously in the text, as, for example, in:

(62) hi1-na1-su2 wa3/lin1-su2-ni3
    today manioc-CL:BONE.LIKE-OBsERVATIONAL.REC.GIVEN
    3-sha1-ra2
    plant-1sg-PERFv
    'Today I planted the manioc roots that we both saw earlier in the day.'

But we left behind a few extra roots.

na1-su1-ai3/na3/k3ai1lu2
    1sg.POSSV-CL;BONE.LIKE-this-therefore
    These my roots therefore,

k?1a1/na1/tai3/na1/na1/te2-y11-7ne2
    tomorrow this-again-MANNER
    3-sha1-tu1-ra2
    plant-1sg-PUT-IMPERFv.though 1sg.POSSV-REC.INTERNAL-PERFv
    I intend tomorrow to plant as before (i.e. as I've done today.)

Note, in the vernacular of this example, the classifier su1 'bone-like' occurs three times (once on each line of the surface forms, where its gloss has been capitalized to facilitate reader recognition). It is easily seen that the second and third occurrences of su1 refer anaphorically to the first occurrence.

However, I do not analyse classifiers as pronouns as they can never occur as pronominal affixes on main verbs. Moreover classifiers can be possessed, but unlike pronouns they can never be possessors. Thus, from the above example, we have:

    1sg.POSSV-CL;BONE.LIKE-this-therefore

where the classifier su1 'bone-like' refers anaphorically to na1/su1/su1/su1 'the manioc roots', first introduced in the first line of the text in the example, and the 1sg.1 '1sg.POSSV' is a possessive pronoun. The whole expression is therefore a possessed nominal, and the classifier -su1 is possessed.

BIBLIOGRAPHY


The Arawá language family was first posited by Brinton (1891: 293) on the basis of vocabularies of Arawá and Paumarí. It consists of six languages, shown in table 11.1. They are located around and between the Purús and Juruá rivers, major southern tributaries of the Amazon. All are in the southern part of the state of Amazonas, Brazil, except for the language with the most speakers, Kulina, which extends to the neighbouring state of Acre and over into Peru. The accompanying map shows the approximate locations of the Arawá languages and dialects. For Arawá, after which the family is named, all we have is about fifty words taken down by the English explorer Chandless in 1867 (these are rather well transcribed); the tribe was wiped out shortly afterwards by an epidemic of measles. The tribes speaking Madi, Dení and Kulina are semi-acculturated – wearing clothes and using guns, salt, sugar, etc. – but still live a fairly traditional life within the jungle, supporting themselves by slash-and-burn agriculture, fishing and hunting. Most members of these tribes speak a little Portuguese (or Spanish, in Peru), as a second language. The Sorowahá were only contacted in 1980; they are monolingual and live a fully traditional life (without clothes or guns). The Paumari have had most intensive contact with non-Indians. Only the older people speak Paumari, the younger ones using just Portuguese or a mixture of Portuguese and Paumari.

Materials available are of uneven quality. Dixon and Vogel have an extensive grammar of Jarawara in an advanced stage of preparation. SIL teams have done some linguistic work on the Jamamadi and Banawá dialects of Madi, on Dení and on Kulina (from the Peru side), producing manuscript sketch grammars for Dení and Kulina and draft dictionaries for all four varieties. For Paumari there is a

\[\text{Map 10 Arawá languages and dialects with approximate locations}\]
Table 11.1 The Arawá language family

<table>
<thead>
<tr>
<th>Language</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paumari</td>
<td>c. 600, only c. 200 speak the language</td>
</tr>
<tr>
<td>Madi</td>
<td></td>
</tr>
<tr>
<td>Sorowahá</td>
<td>c. 100</td>
</tr>
<tr>
<td>Jamamadi</td>
<td>c. 190</td>
</tr>
<tr>
<td>Banawá</td>
<td>c. 80</td>
</tr>
<tr>
<td>Dení-Kulina group</td>
<td></td>
</tr>
<tr>
<td>Dení</td>
<td>c. 1,000</td>
</tr>
<tr>
<td>Kulina (or Madiha or Madija)</td>
<td>2,500</td>
</tr>
<tr>
<td>Arawá (extinct since about 1880)</td>
<td></td>
</tr>
</tbody>
</table>

Table 11.2 Cognate percentages between the Arawá languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Paumari</th>
<th>Madi</th>
<th>Sorowahá</th>
<th>Dení</th>
<th>Kulina</th>
<th>Arawá</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Madi</td>
<td>34</td>
<td>43</td>
<td>40</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
<td>60</td>
<td>61</td>
<td>60</td>
<td>79</td>
<td>53</td>
</tr>
<tr>
<td>43</td>
<td>34</td>
<td>31</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td></td>
<td>76</td>
<td></td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>54</td>
<td>79</td>
<td>53</td>
<td>72</td>
<td>67</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Arawá are based on Chandle's word list and the denominators vary between 19 and 25; these figures have a lower reliability than the others. The three Madi dialects - Jarawara, Jamamadi and Banawá - share about 95 per cent vocabulary, have very similar grammars, and are fully mutually intelligible.

There are enough grammatical differences between Dení and Kulina to establish them as distinct languages, but they plainly constitute one subgroup. No other subgrouping is justified on the information currently available. Madi has strong lexical and grammatical similarities with Dení and Kulina whereas Paumari is more divergent. As mentioned in chapter 1, there must have been a great deal of tribal merging during the past few hundred years. The Paumari and Sorowahá tribes may be the result of such mergers, one or more tribes speaking an Arawá language coming together with a tribe or tribes speaking languages from other families; this would have resulted in a significant substratum vocabulary of non-Arawá origin.

### 2 Phonology

There is a straightforward (C)V syllable pattern in all languages. For proto-Arawá, I have provisionally reconstructed the consonant system set out in table 11.3 (based on Dixon, forthcoming-b). Chapman and Derbyshire (1991: 347) mention that in Paumari there is a contrast between a flap vibrant [f] and a retroflexed grooved reverse flap vibrant [f] (although this has a very low functional load). These two sounds may have been present in proto-Arawá. All other languages have a single liquid [r], usually with both rhotic and lateral allophones.

Paumari has the most conservative system. Both *ts and *ts s have become s; *t has become a glottal stop medially but is retained initially; *ph has been replaced by f, and *dz has become dz. The Paumari vocabulary includes a substratum (which may come from another Arawá language) where *p > ?, *dz > s, *p > h and *s > h (see Dixon 1995: 291–3).

Dení, Kulina and Madi have lost the imploded voiced stops, with *b > b initially and >p medially, and *d > t in all positions. They have also lost g (generally replaced by w) and most or all occurrences of the glottal stop. The only other change in Dení and Kulina is *f > tf.

Madi has undergone further changes. First, aspiration was lost, with *p > p,
The consonant system of the Jamamadi and Banawá dialects of Madi

<table>
<thead>
<tr>
<th>voiced stop</th>
<th>bilabial</th>
<th>apico-dental</th>
<th>apico-alveolar</th>
<th>lamino-palatal</th>
<th>dorso-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>J</td>
<td>k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless stop</td>
<td>t</td>
<td>m</td>
<td>n</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>θ</td>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>semi-vowel</td>
<td>(w)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11.5 Proto-Arawá vowels

<table>
<thead>
<tr>
<th>front</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>o</td>
</tr>
<tr>
<td>mid</td>
<td>e</td>
</tr>
<tr>
<td>low</td>
<td>a</td>
</tr>
</tbody>
</table>

Proto-Arawá had a system of four vowels, given in table 11.5. These are maintained in Dení, Kulina and Madi. In Paumari and Sorowahá, e has been lost, generally being replaced by i in the third syllable of a root and by a elsewhere.

It is unusual in Amazonia to find a vowel system that lacks a high central vowel /i/. In fact, younger speakers of Sorowahá have innovated [i] as an allophone of any of the three vowels /i/, /a/, and /o/ – in unstressed position.

In the three languages in which e is retained, it plays a pervasive role in engendering assimilation of a preceding a, e.g., *ama-ne ‘blood-MASC’ has become eme-ne in Dení, Kulina and Madi (it is ana-ne in Paumari).

There is an interesting link between semivowel w and vowel o (this was first pointed out in Adams Lichlan and Marlett 1990). There is no contrast between o- and vo- at the beginning of a word, or between -owV- and -oV- or -Vwo- and -Vo- (where V is a vowel other than o) medially, and there is no example of -owl-vo-. However, we do get a contrast between V- and wV- initially (e.g., *adami ‘hill’ and *wadami ‘dream’) and between -VwV- and -VV- medially. It is possible to treat w and o as allophones of a single phoneme (which we can call /O/) but the cost of this ‘economy’ is that a syllable boundary must always be marked. Thus /Oinaw/ could be either oinawina or wina, which are different words. The first would have to be shown as /O.ina/ and the second as /O.ina/, where ‘.’ indicates a syllable boundary.

One characteristic of Arawá languages is a profusion of lexical homonymy, in which speakers appear to delight. The neutralizations that have taken place (e.g., *d, *l, *t, and *l all falling together as t in Jarawara) have added to this number. For instance, in Jarawara there are three nouns jikjo: (a) ‘hammock’ (masc.); (b) ‘fire, firewood’ (fem.); and (c) ‘miriti palm (Mauritia veifera)’ (masc.). And there are four verbs saa: (a) ‘shoot fish with arrow’; (b) ‘release, let go’; (c) ‘peel, strip’; and (d) ‘vomit’.

Arawá languages are also prone to referential merging, as illustrated in table 11.6 for the words for ‘moon’, ‘sun’ and ‘thunder’. It will be seen that the Jamamadi dialect of Madi maintains the original maha ‘sun’ and bahi ‘thunder’. In Jarawara, bahi ‘thunder’ has been extended also to cover ‘sun’ while in Banawa maha ‘sun’ has been extended also to cover ‘thunder’. In Sorowahá, masiki ‘moon’ has been extended also to cover ‘sun’. (Paumari appears to have innovated safini for ‘sun’, but it does have a lexeme maha meaning ‘day’ or ‘time’ as a possessed noun, ‘life’, as in ‘my life’.)
3 WORD CLASSES

Arawá languages have two open lexical classes - nouns and verbs. They typically have a small closed class of adjectives, which can only modify a noun within an NP. A class of about fourteen adjectives is given for Dení and one of similar size for Jarawara. These include reflexes of *'boడe 'old' and *da'ɖi 'new' and also terms for 'big', 'small', 'other'. There are closed classes of pronouns, demonstratives, postpositions, clause linkers and interjections.

4 NOUNS

These divide into two subclasses - free and possessed. Each free noun has an inherent gender, marked not in the form of the noun but in the form of modifiers within an NP and suffixes to the verb. There are two genders - feminine, which is the functionally unmarked term (for instance, all pronouns are cross-referenced as feminine), and masculine.

A possessed noun follows a free noun or pronoun and agrees with it in gender. There can be well over 100 possessed nouns, including body parts, parts of trees, and such things as 'na me', 'home', 'dream', 'food', 'path' and 'container'. There are also possessed nouns referring to orientation, e.g. 'in front of', 'on top of', 'inside', 'by the side of'; these correspond semantically to prepositions or postpositions in other languages.

In proto-Arawá, gender was marked on nouns by the suffix -ni for feminine and -ne for masculine. All languages retain feminine -ni, except Madi where it has been lost, often leaving a phonological trace, e.g. *noko-nti > noki 'eye + FEM', *mata-nti > mate 'buttock, tree stump + FEM'. In all languages masculine -ne has been lost from some, but not all, possessed nouns. In Dení, Kulina and Madi it has often engendered vowel assimilation before being lost, e.g. *mata-ne > mete-ne > mete 'buttock, tree stump + MASC'. As a result of these changes, in Madi about half the possessed nouns have identical feminine and, masculine forms, e.g. *'dd'ndi-ne > tati/fati 'head', in all other languages the gender distinction is maintained, through the retention of suffix -ni. (There is a full discussion of possessed nouns in Dixon 1995.)

Just in Paumarí, Dení and Kulina there is a further division of nouns, into those that require a cross-referencing prefix ka- on the verb and on some nominal modifiers (when the noun is in pivot function in the clause) and those that do not. A relatively small number of nouns belong to the ka- class; they include some possessed nouns, and free nouns of a number of semantic types. There is a partial semantic basis - containers and things in containers (e.g. 'pan', 'canoe') and things made up of parts (e.g. 'house', 'plane') or of particles ('sugar', 'banana mash') are generally ka-class. Some nouns can be ka- or non-ka-, with a meaning difference, e.g. in Paumari sa'ni is 'hand' as ka-class and 'finger' as non-ka-, while siho is 'fire' as ka-and 'firewood' as non-ka-class.

5 NOUN PHRASE STRUCTURE

The general structure of a noun phrase is:

(a) alienable possessor. This can be a full noun phrase, and is followed by possessive marker *kha.
(b) noun or pronoun. This supplies the gender - and the ka-specification - for the whole noun phrase.
(c) adjective(s) - one or more members of the closed class (some show distinct gender forms).
(d) possessed noun(s), agreeing in gender with the head noun. There may also be a plural marker, e.g. deni in both Dení and Kulina.

This structure can be illustrated by a sample noun phrase in Jarawara. Note that the noun jomee 'dog', is both alienable possessed and inalienable possessor:

(1) [fana kaa jomee bote tene
woman(fem) possv dog(masc) old foot + masc
'the woman's old dog's foot'

Within an NP showing alienable possession it is the possessed noun which is head. However, the question of what is the head of an NP which consists of a pronominal possessor and a possessed noun is not a simple matter. An NP such as o-mano 'my lower arm' triggers feminine agreement on the verb, relating to the 1sg o- (all pronouns are cross-referenced with the unmarked gender, feminine). However, the NP counts as 3rd person on other criteria (e.g. in an O-construction where the verb will show prefix hi- if both A and O arguments are 3rd person - see §10 below). That is, from the point of view of gender agreement the pronominal possessor appears to be the head, but from the point of view of person agreement it is the possessed noun that appears to be the head.

 Demonstratives are typically apposed to noun phrases (rather than acting as direct modifier of a head noun). There is a limited set of enclitics or postpositions that may follow a noun phrase, indicating its function in the clause. These include:
(i) accusative -ra can follow a noun phrase in O function in Paumarí, in Sorowahá and in the Jamamadi dialect of Madi (in Jamamadi it may only attach to an O NP if the A argument is third person). In the Jarawara dialect it is now restricted to occurrence on pronouns.

(ii) just Paumarí has enclitic -a with a locative and instrumental meaning. It is also used to mark a noun phrase in A function in an O-construction (see below).

(iii) in Paumarí, *dza has directional meaning 'to'. In Dení, Kulína and Madi it has a much wider function, covering some or all of locative ('at'), allative ('to'), ablative ('from'), comitative and instrumental ('with').

6 VERBS

It is likely that in proto-Arawá most (or all) verbs could function in either of two ways (with a meaning difference) — either accepting prefixes and suffixes themselves, or having prefixes and suffixes added to a following auxiliary, -na- or -hi- (the forms are -ni- and -hi- in Paumarí). This is retained to some extent in Paumarí, e.g. we getmitha- (with affixes on the verb)'hear' and mitha-nti- (with affixes on the auxiliary) 'listen'. In some languages verbs have separated out into two subclasses — inflecting (taking affixes themselves) and non-inflecting (requiring an auxiliary to carry the affixes). There is a partial semantic basis — for instance, most stative verbs (corresponding to adjectives in other languages) are inflecting, e.g. 'be hard', 'be cold', 'be black', 'be sweet', 'be angry'.

There is also a division of verbs into intransitive (e.g. 'sit'), transitive (e.g. 'hit'), amitransitive where S of the intransitive corresponds to O of the transitive (e.g. 'break'), and amitransitive where S corresponds to A (e.g. 'paddle (a canoe)'). There is no correlation between transitivity and the inflecting/non-inflecting division.

In Madi and Kulina (at least) there are verbs with suppletive forms depending on whether the S or O argument is singular, dual or plural; these include 'sit', 'lie', 'put', 'fall', 'hit' and 'kill'.

7 PREDICATE STRUCTURE

The only obligatory constituent of a clause is the predicate. Its structure is typically:

(a) object pronoun (a separate word); obligatory in a transitive clause.

(b) subject pronoun (may be a prefix or a separate word); obligatory in all clauses.

(c) prefix ka-; this can have a number of functions: (i) cross-referencing a ka-class noun in pivot argument function; (ii) marking an applicative derivation, in which a transitive stem is derived from an intransitive root, with underlying S becoming A and a peripheral argument being promoted to be O, e.g. 'laugh + at'; (iii) a variety of other semantic functions, e.g. dual subject, 'accompanied by', 'inside'.

(d) causative prefix na- or niha-; in all languages this derives a transitive stem from an intransitive root, underlying S becoming surface O. In some it can also be used with transitive roots, the underlying A becoming surface O and the underlying O going into a peripheral slot. In Paumarí a transitive verb must first be detransitivized (by suffix -a) before taking causative prefix na-.

(e) verb root, inflecting or non-inflecting.

(f) auxiliary verb -nali- or -hua- (there are other, minor, possibilities in Dení and Kulina); obligatory with a non-inflecting verb root, taking prefixes (b)-(d) and all suffixes.

Arawá languages have rich suffixal possibilities — several score forms (organized into fifteen or more slots), but all are optional. The major types are:

(g) a large class referring to location ('on the ground', 'in the water'), direction ('up', 'alongside'), duration ('all night long', 'continuously'), and 'following', 'almost', 'again', etc; the same sorts of meaning are expressed by verbal suffixes across languages of the family but the actual forms tend to differ. Just a few forms recur, e.g. *-thima 'upriver'.

(h) a negative suffix; the Arawá word list shows a negative suffix -raha. Kulina has two negators: static verbs take -ra immediately after the root or auxiliary while non-statics have -hera(temyl)-hara(masc) after tense-aspect. Dení just has -hera-hara. Jarawara has -ra which comes either before tense-aspect or after mood (when there is no tense-aspect suffix). Paumarí and the other dialects of Madi have -ra, before tense-aspect. (In Sorowahá there is a negative suffix with quite different form, nasiol) In Madi (and probably in some other languages) the only way of expressing negation is through a verb plus negative suffix; there is no interjection 'no'.

(i) tense-aspect suffixes; all languages have a set of suffixes for indicating reference to time, but there is considerable semantic variation. Chapman and Derbyshire (1991: 275ff.) state that there is no specification of tense as such in Paumarí, but there are aspectual suffixes 'comitative', 'continuous', 'iterative', 'durateive', etc. At the opposite extreme, Jarawara
The main pronominal forms are shown in Table 11.7. Those in row (i) function as subject of a verb (in slot (b) of predicate structure) and as possessor of a possessed noun. Those in row (ii) are the free forms, also generally functioning as object (in slot (a) of the predicate), sometimes with the addition of accusative suffix -ra. Paumari has nothing corresponding to the set (ii) of other languages; here a further set of prefixes (1sg -ho, 2sg -i, 1non-sg hari, 2non-sg aro) can be added to one of a small number of forms (e.g. -ra, accusative, -wani, contrastive).

Table 11.7 Main pronominal forms in Arawá languages

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
<th>1non-sg</th>
<th>2non-sg</th>
<th>3non-sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paumari (i)</td>
<td>o-</td>
<td>i-</td>
<td>ø</td>
<td>a(r)a-</td>
<td>awa-</td>
<td>wa-</td>
</tr>
<tr>
<td>Dení (i)</td>
<td>o-</td>
<td>ti-</td>
<td>ø</td>
<td>i-</td>
<td>ti- + redupl.</td>
<td>ø</td>
</tr>
<tr>
<td>Kulina (i)</td>
<td>o-</td>
<td>ti-</td>
<td>ø</td>
<td>i-</td>
<td>ti deni</td>
<td>poni deni (fem)</td>
</tr>
<tr>
<td>(ii)</td>
<td>owa tiwa</td>
<td>poni (fem)</td>
<td>poni deni (fem)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poa (masc)</td>
<td>poa deni (masc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jamamadi (i)</td>
<td>o-</td>
<td>ti-</td>
<td>ø</td>
<td>i-</td>
<td>ti ari</td>
<td>poni deni (fem)</td>
</tr>
<tr>
<td>(ii)</td>
<td>owa tiwa</td>
<td>hine</td>
<td>poni deni (fem)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorowahá (i)</td>
<td>o-</td>
<td>i-</td>
<td>ø</td>
<td>ari</td>
<td>ari</td>
<td>ari</td>
</tr>
<tr>
<td>(ii)</td>
<td>aro taa</td>
<td>ini, iri</td>
<td>ari</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Alienable possession is generally shown by a prefix of set (i) added to possessive marker *klu'. Again, Paumari is different, with forms such as 1sg kodi (possibly < o-kadi), 1non-sg a-kadi.

Only in the Madi dialects is there a distinction between inclusive and exclusive in 1non-sg. The inclusive form is e, corresponding to general 1non-sg in other languages, and oda (otaa in Jarawara) has been innovated as the exclusive pronoun. 3sg is always zero in set (i). The 3sg set (ii) forms in Dení and Kulina (poni and po(w)a) are used for emphasis. The 3sg set (ii) forms in the Madi dialects (hini in Jamamadi and hinihina in Jarawara) occur only in reflexives. The Madi dialects (and probably also Sorowahá) have developed a 3non-sg pronoun — in sets (i) and (ii) — from the original noun *madi 'person'; we can observe changes *madi > mai > mee. Just Paumari has a 3non-sg prefix, wa-.

Commenting now on the 1st and 2nd person forms:

(a) 1sg shows o- for verbal and possessive prefix, in all languages. (The 1869 Arawá word list includes 'usafá, hand' and 'otarná, foot', which plainly include prefix o- 'my'.) The set (ii) form is owa in Dení, Kulina and Madi and aro in Sorowahá.

8 PRONOUNS

The main pronominal forms are shown in table 11.7. Those in row (i) function as subject of a verb (in slot (b) of predicate structure) and as possessor of a possessed noun. Those in row (ii) are the free forms, also generally functioning as object (in slot (a) of the predicate).
(b) 2sg set (i) was plainly prefix *ti- in proto-Arawá, with the t being lost in Paumari and Sorowahá. The set (ii) form is *tiwa in Madi and Kulina, *tia in Dení and *tao in Sorowahá. (Paumari has a couple of traces of an archaic ti, e.g. in the greeting *iwan(l)i which involves the 2sg free pronoun *iwa plus an optional -ti.)

In Sorowahá 1sg o- and 2sg i- are prefixes to possessed nouns but inflxes on a verb beginning with a consonant, following this consonant. Thus, from gana- 'see', we get g-o-ania 'I see', and from sawa- 'wash', there is s-i-awa 'you wash'. In fact the infix is phonetically an off-glide, i.e. [g'ania] and [s'awa] respectively.

(c) For 1non-sg, *ari occurs as the set (ii) form in Dení and Sorowahá and as set (i) prefix in Paumari. Just Dení and Kulina have a set (i) prefix i-; it is not clear whether this should be reconstructed for proto-Arawá or whether it is an innovation in Dení-Kulina (perhaps as the reduced form of an earlier prefix ar-i). In Madi and Sorowahá the same free form is used in both sets. It seems likely that *ari reduced to ai in Jamamá and then to e in Jarawara. The Kulina form is may also relate to *ari (or there may be some other explanation for this).

(d) 2non-sg shows the greatest variation. The Madi dialects have free form deeo (tee in Jarawara) in both set (i) and set (ii) and Sorowahá has *tima in set (ii). Paumari shows a prefix awa-. Kulina has the same prefix, ti-, as 2sg and adds a plural suffix -deni. Dení also uses prefix it- but here plural is shown by reduplication; roughly, if a verb (omitting the final mood suffix) ends in \(-CV\) or \(-CV\#\), then \(V^p\) is repeated before \(C_i\), giving \(-V^pCV_iV^p\#\), e.g. ti-kha-thima-aro (‘2-be in motion-UPSTREAM-DECL + FEM’) ‘you (sg) go upstream’ and ti-kha-thi-q-ma-aro ‘you (pl) go upstream’.

9 DEMONSTRATIVES AND INTERROGATIVES

There is typically a rich set of demonstratives, but the semantic details and forms vary from language to language.

Most Arawá languages have a single root underlying most or all of 'who', 'what', 'where', etc. This is nahiha (with variants niha and nana) in Paumari, neheko in Kulina and ako in Dení; in Dení and Kulina feminine -ha)ro and masculine -ha)ri are added to the basic root to form 'who'. In Jamamá 'who' is ebenike and in Jarawara it is hika (fem) and higha (masc) which may be further cognates.

10 CONSTRUCTION TYPES

There are two transitive construction types, the choice between them depending on whether the A or the O argument is pivot (topic) within the discourse in which the clause appears. See examples (2–3) below.

The main characteristics of the two construction types are:

A-CONSTRUCTION

The A argument is discourse pivot
The A NP is generally not included
Verbal suffixes agree with A
O marked by accusative -ra

O-CONSTRUCTION

The O argument is discourse pivot
The O NP is generally not included
Verbal suffixes agree with O
A marked by ergative -a (only in Paumari)

The two construction types have different preferences for constituent order. In the few textual instances where both A and O NPs are stated, Madi prefers AOV for an A-construction and OAV for an O-construction while Paumari prefers OVA for an A-construction and AVO for an O-construction. (That is, the pivot NP is preferred in clause-initial position in Madi and immediately after the predicate in Paumari.)

The A-construction appears to be the unmarked construction type, and can be used for all possibilities at A and O. The O-construction is more restricted – in Paumari it may only be used if the O is 3rd person, in Dení and Kulina only if both A and O are 3rd person, and in Jarawara only if either A or O (or both) is 3rd person.

Sentence (2) gives an example of an A-construction and (3) of an O-construction from Jarawara. Each is preceded by an intransitive clause to provide discourse context.

(2) fana, to-ka-hara-ke
woman(fem) away-be.in.motion-IMM.PAST.eyewitness + fem-DECL. + fem

awiy o-mita-hara-ke
tapir(masc) 3sg-hear-IMM.PAST.eyewitness + fem-DECL. + fem

‘The woman went out [into the forest] and heard a tapir.’

(3) awiy ka-ke-hare-ka
tapir(masc) be.in.motion-COMING-IMM.PAST.eyewitness + masc-DECL. + masc
fana, hi-mita-re-ka
woman(fem) O-construction-hear-IMM.PAST.eyewitness + masc-DECL. + masc

‘A tapir came [into the village] and the woman heard him.’
In (2), *fana* 'woman' is the pivot running through the two clauses. The transitive clause is then an A-construction in which *fana* is the understood A and *awi* 'tapir' the stated O. It is recognizable as an A-construction by the absence of a verbal prefix *hi*- and by the feminine forms of tense and mood suffixes, agreeing with the A argument 'woman'. In (3), *awi* 'tapir' is the pivot linking the two clauses; it is in S function in the first and in O function in the second. The latter is an O-construction, marked by verbal prefix *hi*- and the masculine forms of tense and mood suffixes, agreeing with the O argument. (A fuller discussion of A-constructions and O-constructions in Jarawara is in Dixon forthcoming-a.)

Arawá languages also have complement clauses (filling a core slot in the main clause) and a rich array of other types of subordinate clause constructions. Just Paumarí also has a passive derivation (in addition to A-constructions and O-constructions).

**BIBLIOGRAPHY**

Adams, P. 1987. 'Madíja grammar sketch' (87p typescript, SIL Peru).
Dixon, R. M. W. forthcoming-b. 'Proto-Arawá phonology'.

---

12

Small language families and isolates in Peru

MARY RUTH WISE

The members of five small language families — Cahuapana, Jivaró, Peba-Yagua and Witoto — are located in northern Peru and across the borders in Ecuador, Colombia, and Brazil. The speakers of these languages, with the exception of Jivaró and possibly Cahuapana, were among the ethnic groups whose populations were decimated during the 'rubber boom' at the turn of the twentieth century. The indigenous population of the Putumayo River region dropped from 50,000 to 7,000-10,000 during the first decade of the twentieth century. The Zaparo languages have lost ground continually to Quechua; they and all of the families except Jivaró are also losing ground to Spanish.

One other small language family might be added to the five discussed in this chapter, namely Huraknubit (or Hate) of the southern Peruvian jungle. However, the current consensus is that it should be considered a single language isolate. Two isolates are almost extinct. Language loss appears to have occurred in Munich in the form of morphological simplification (Gibson 1996: 26). It was impossible to record any folklore among the Taushiro since all of the old story tellers died in an epidemic in about 1964 (N. Alicea p.c.).

Table 12.1 lists the languages within families or as isolates, as well as other languages of northern Peruvian Amazonia; alternative names within parentheses are followed by location, population, and comments on literacy, bilingualism and dialects. Unless specified otherwise, references to bilingualism indicate proficiency in the vernacular and Spanish.

---

1 Unless specified otherwise, references to bilingualism indicate proficiency in the vernacular and Spanish.
Table 12.1 Small language families and isolates in Peru

<table>
<thead>
<tr>
<th>Family</th>
<th>Subfamily</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cahuapana</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a Chayahuita</td>
<td>(Canpo piyapi, Tshaawi)</td>
<td>Paranapura, Cahuapana, Sillay and Shanusi rivers; pop. 7,000; 20% incipient to quite bilingual; 30% literate in Chayahuita and Spanish; dialects: Chayahuita, Cahuapana.</td>
</tr>
<tr>
<td>1b Jebero</td>
<td>(Chebero, Xiwila)</td>
<td>District of Jeberos; pop. 2,000; only older adults speak the language.</td>
</tr>
<tr>
<td>Jivaro (Shuar)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a Aguaruna (Aguyajun)</td>
<td></td>
<td>Western Upper Marañon, Potro, Mayo and Cahuapanas rivers; pop. 39,000; 35% monolingual; at least 80% literate in Aguaruna and many (65%) in Spanish also.</td>
</tr>
<tr>
<td>Shuar subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2b Achuar-Shiwiar (Achual, Jivaro, Maina)</td>
<td></td>
<td>Morona, Macusari, Tigre, Huasaga, and Corrientes rivers in Peru; Pastaza and Bobonaza rivers in Ecuador; pop. 5,500; 90% monolingual; 35% literate.</td>
</tr>
<tr>
<td>2c Wambisa</td>
<td></td>
<td>Morona and Santiago rivers; pop. 6,000–10,000; 20% monolingual; 50% literate.</td>
</tr>
<tr>
<td>2d Shuar (Jivaro, Chiwara, Xivar)</td>
<td></td>
<td>Morona-Santiago Province in Ecuador; pop. 32,000; 14% monolingual; 50% literate.</td>
</tr>
<tr>
<td>Zaparo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a Andoa (Shimigae, Semigae, Gae, Gaye)</td>
<td></td>
<td>Village of Andoa, Pastaza river, pop. of ethnic group – 150–260, language became extinct in 1993; all now speak Southern Pastaza Quechua and/or Spanish.</td>
</tr>
<tr>
<td>3b Arabela (Chiripuno)</td>
<td></td>
<td>Arabela river, tributary of Napo; pop. of ethnic group – 300, speakers – 150 maximum; no monolinguals in Arabela, many trilingual in Arabela, Tigre (Northern Pastaza) Quechua and Spanish. Note, however, that a small isolated subgroup (the Pananahurt) which separated from those on the Arabela about 100 years ago may still exist.</td>
</tr>
<tr>
<td>3c Aushiri (Auxira)</td>
<td></td>
<td>Tributaries of right bank of Napo river; extinct.</td>
</tr>
<tr>
<td>3d Cavaran</td>
<td></td>
<td>Headwaters of the Nanay river; 5 speakers in 1972; possibly a dialect of Iquito.</td>
</tr>
<tr>
<td>3e Iquito (Atacacore, Hamacore, Quiturcan, Puca-Uma)</td>
<td></td>
<td>Northern Nanay river; pop. 150 in 1965; only older adults speak the language, all are bilingual with Spanish.</td>
</tr>
<tr>
<td>3f Omurano (Roamaina, Mayna)</td>
<td></td>
<td>Urituyacu river; extinct by 1958.</td>
</tr>
<tr>
<td>Table 12.1 (cont.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>3g</strong> Zaparo (Zapara, Kayapwe)</td>
<td>Pastaza Province, Ecuador; pop. of ethnic group – 150, only 2/3 older adults spoke the language in 1980; integrated with the Northern Pastaza Quechua.</td>
<td></td>
</tr>
<tr>
<td><strong>Peba-Yagua</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4a</strong> Yameo</td>
<td>Itaya and Nanay rivers; pop. 50 in 1925; now extinct.</td>
<td></td>
</tr>
<tr>
<td><strong>4b</strong> Peba</td>
<td>formerly spoken north of the town of Pebas on the Amazon river; extinct.</td>
<td></td>
</tr>
<tr>
<td><strong>4c</strong> Yagua (Yawa, Llugua, Yegua)</td>
<td>Northeastern Amazon river from Iquitos to Brazil border, a few in Colombia and Brazil; pop. 3,000–4,000; 25–35% bilingual in Spanish, low level of literacy; some dialectal differences.</td>
<td></td>
</tr>
<tr>
<td><strong>Witoto</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bora–Muínane subgroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5a</strong> Bora</td>
<td>Colombia–Peru border, mostly in Peru – Yaguasuyacu, Ampiyacu and Putumayo rivers; pop. 2,500–2,500; high rate of literacy and bilingualism; 94% mutually intelligible with Miraña dialect (100 speakers in Brazil and Colombia).</td>
<td></td>
</tr>
<tr>
<td><strong>5b</strong> Muínane</td>
<td>East Central Amazonas in Colombia; pop. 150; all are bilingual with Bora or Witoto.</td>
<td></td>
</tr>
<tr>
<td><strong>Witoto–Ocaina subgroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5c</strong> Ocaina</td>
<td>Yaguasuyacu and Ampiyacu rivers, mostly in Peru, a few in Colombia; pop. 150–200; multilingual in Murui Witoto, Bora and/or Spanish; dialects: Dukaiya, Ibo’te.</td>
<td></td>
</tr>
<tr>
<td><strong>Early Witoto languages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5d</strong> Nipode (Muínane) Witoto</td>
<td>mostly in Peru; pop. 50–100; bilingual in Murui Witoto and/or Spanish.</td>
<td></td>
</tr>
<tr>
<td><strong>5e</strong> Minica Witoto</td>
<td>Igaraparana river in Colombia; pop. 2,500; many are bilingual with Spanish and literate.</td>
<td></td>
</tr>
<tr>
<td><strong>5f</strong> Murui Witoto (Báe)</td>
<td>Putumayo, Ampiyacu, and Putumayo rivers; pop. 2,000–2,800; 95% of those under 40 are literate; 90% of those under 50 are at least incipient bilinguals.</td>
<td></td>
</tr>
<tr>
<td><strong>Isolates (includes only extant languages)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>6</strong> Candoshi-Shapra (Murato)</td>
<td>Huallaga River Valley; 1 or 2 speakers in 1985; no recent data available.</td>
<td></td>
</tr>
<tr>
<td><strong>7</strong> Cholen (Tinganeses, Seeptsa)</td>
<td>Madre de Dios, Colorodo and Keros rivers; pop. approx. 650; approx. 80% are at least incipient bilinguals; approx. 10% literate in Amarakaeri and 50% in Spanish; major dialects: Amarakaeri and Wachipaeri; Kisambaeri is probably a subdialect of the former and Amsaari, Sapiteri and Toperi are probably subdialects of the latter.</td>
<td></td>
</tr>
<tr>
<td><strong>8</strong> Harakmbet (Hate)</td>
<td>village of Munichis on the Prawnpara river; 3 speakers in 1988.</td>
<td></td>
</tr>
<tr>
<td><strong>9</strong> Munichi (Munichino, Otanabe)</td>
<td>Aucayacu river, a tributary of the Tigre; pop. approx. 5.</td>
<td></td>
</tr>
<tr>
<td><strong>10</strong> Taushiro (Pinche)</td>
<td>an isolate spoken along the Amazon in Peru, Colombia and Brazil; pop. 6,000 in Peru, 14,000 in Brazil, 5,000 in Colombia; high rate of literacy in Spanish and Ticuna among Peruvian Ticuna; many bilingual with Spanish or Portuguese.</td>
<td></td>
</tr>
<tr>
<td><strong>11</strong> Ticuna (Tukuna)</td>
<td>Chambira and Uruyaysucu rivers; pop. 3,500; 40% at least incipient bilinguals; some literate in Uruyana and Spanish.</td>
<td></td>
</tr>
</tbody>
</table>
related, are mutually unintelligible. No reconstruction has been proposed, nor has membership in a larger family or stock been confirmed by comparative work.

The Jivaro (Shuar) family includes Aguaruna (the most diverse), Wambisa, Achuar and Shuar. Candoshi has sometimes been classified as a member of a Shuar–Candoshi family but David Payne (1989) considers his earlier (1981) comparative reconstruction of proto-Shuar-Candoshi to be unjustified. It has also been erroneously classified as Zaparo by some authors.²

Little data are available on Peba and Yameo and no comparative work has been done on the Peba-Yagua family.³

Languages of the Zaparo family are spoken in the eastern jungle of northern Peru and southeast Ecuador. The whole family is nearly extinct with fewer than 200 speakers in total. Only Arabela and Iquito are spoken by more than 10 people, and only Arabela has even a slim chance of survival beyond the current generation of older adults. (See Stark 1990 for a description of the sociolinguistic situation.) Taushiro is possibly a Zaparo language, but its classification has not been confirmed.

The Witoto family includes two main branches: Bora–Muinane and Witoto–Occina. The two Bora–Muinane languages are: Bora, including the Miraña dialect which is about 94% intelligible with Bora, and Muinane, not to be confused with Muinane Witoto. Resigaro has sometimes been included in the family but David Payne (1985) demonstrates that the apparent relationship is the result of heavy borrowing, with consequent phonological change and some grammatical borrowing; Resigaro is clearly an Arawak language (see also Aschmann 1993: 3).

Harakmbet has sometimes been erroneously classified as Pre-andine Arawak. Apparent cognates are probably due to borrowing and to the presence of some widespread grammatical forms (cf. David Payne 1990b). There has been some convergence of dialects, or rather extension of Amarakerí, at least in the Dominican Mission at Shintuya (Van den Eynde 1972: 4; see also Lyon 1975).

2 PHONOLGY

Most of the languages in the five families differ from areal patterns in one or more traits. In the Jivaro languages there are few restrictions on syllable-final consonants so that consonant clusters are very frequent. Syllables are usually open in the other languages. Aguaruna (Jivaro) has no liquids while Jebero (Cahuapana) has three. Both fricatives and affricatives occur in some but others have only one or the other.

A four-vowel system is the most frequent but several languages have six.

The consonant and vowel inventories of Jebero (Cahuapana) are shown in tables 12.2 and 12.3 respectively (adapted from Bendor-Samuel 1961).

Chayahuita lacks c and the glottalized consonants, has only one liquid (r) and two nasals (bilabial and alveolar). In Jebero, occlusives are voiced after nasals. The phoneme /s/ is ‘produced by the blade of the tongue in the palatal region with the tongue tip down behind the bottom teeth’ (Bendor-Samuel 1961: 13); s is a frictionless continuant. Syllables can be open or closed. The central near-close vowel is lower in Jebero than in Chayahuita. 7, r-quality and nasalization are analysed as syllable prosodies in Jebero (1961: 20). Vowel length, aspiration, glottalization [vowel + ?] and nasalization are considered to be part of the syllable nucleus in Chayahuita (Hart and Hart 1976) so that syllables are open. Stress is not contrastive in either language. In Jebero the first syllable of a disyllabic word is accent and the second of words with three or more syllables; there is some modification with certain suffixes.

David Payne (1989) suggests a possible deep genetic relationship between Candoshi and Arawak; he presents thirty-two basic vocabulary forms with CVC resemblances and several apparent grammatical cognates. Gnerre (1988) suggests a possible deep genetic relationship between the Jivaro and Arawak families. Neither suggests that Jivaro languages and Candoshi are more closely related to each other than to Arawak.


² David Payne (1989) suggests a possible deep genetic relationship between Candoshi and Arawak; he presents thirty-two basic vocabulary forms with CVC resemblances and several apparent grammatical cognates. Gnerre (1988) suggests a possible deep genetic relationship between the Jivaro and Arawak families. Neither suggests that Jivaro languages and Candoshi are more closely related to each other than to Arawak.
Table 12.4 Aguaruna consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>alveolar</th>
<th>post-alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>labio-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>f</td>
<td>g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>ts</td>
<td>tf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>s</td>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>approximant</td>
<td>j</td>
<td>v</td>
<td>w</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12.5 Aguaruna vowels

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>close</td>
<td>i</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>open</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Jivaro languages are characterized by many consonant clusters word-medially. The consonants and vowels of Aguaruna are shown in tables 12.4 and 12.5 (adapted from David Payne 1981).

Nasal and oral vowels contrast. As in many Amazonian languages, nasalization can extend over a whole series of vowels and semivowels; nasal consonants are pronounced as prenasalized voiced plosives in the environment of oral vowels. (See David Payne 1976 for a detailed analysis of nasality; David Payne 1990a) proposes that Aguaruna has a type of pitch accent in which stress (intensity) does not necessarily coincide with high pitch (accent) as shown in (1). This analysis is probably applicable to the other Jivaro languages. The reflex of */ŋ/ is /ŋ/ in syllable-initial position (often syllable-final after vowel deletion) in all of the languages except Aguaruna which retains /ŋ/, consequently there are no liquids in Aguaruna. Voiceless word-final vowels are contrastive but are no longer pronounced by many speakers. Contrastive vowel length is interpreted as a vowel sequence. Syllables may be open or closed (2a). Vowel deletion and/or metathesis result in many consonant clusters (2b).

(1) (David Payne 1990a: 165)

[Tʃu]/[tʃu/ 'monkey (Nom)' vs [ʃu]/[ʃu/ 'monkey (ACC)'
(David Payne 1981: 336)

[kap]/[kaap/ 'vine' vs [kaap]/[kaap/ 'gnat'

4 Throughout this chapter the acute accent (') is used to indicate high pitch or tone. Low tone is left unmarked in those languages where tone is contrastive.

Table 12.6 Yagua consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>alveolar</th>
<th>post-alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>labio-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>p</td>
<td>[p, p]</td>
<td>t</td>
<td>[t]</td>
<td>k</td>
<td>[k]</td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>[m, m]</td>
<td>n</td>
<td>[n, n]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>ts</td>
<td>[ts, ts]</td>
<td>f</td>
<td>[f]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>s</td>
<td>[s]</td>
<td></td>
<td></td>
<td></td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>r</td>
<td>[r]</td>
<td></td>
<td></td>
<td>j</td>
<td>[j]</td>
<td>w</td>
</tr>
</tbody>
</table>

Table 12.7 Yagua vowels

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>close</td>
<td>i</td>
<td>[i, i]</td>
<td>u</td>
</tr>
<tr>
<td>open</td>
<td>a</td>
<td>[a]</td>
<td></td>
</tr>
<tr>
<td>close-mid</td>
<td>e</td>
<td>[e]</td>
<td></td>
</tr>
<tr>
<td>open-mid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>open</td>
<td>u</td>
<td>[a, a]</td>
<td></td>
</tr>
</tbody>
</table>

(2a) (David Payne 1981)

/nanap/ 'meat'
/suwatf/ 'lungs'
/pigkín/ 'good'

(2b) (Wipio et al. 1996)

/saunánk/ 'kind of leaf' vs /saunkántín/ 'stream where plants with these leaves abound'
/fínunik/ 'to draw near' vs /funúmtut/ 'to draw near to someone'

The phonemes of Yagua and their principal allophones except for most palatalized consonants are shown in tables 12.6 and 12.7 (adapted from Doris Payne 1985b: 16–17).

Powlison (1995: 30–3) summarizes the effects of /ʃ/ on vowels: for example, /wʃa/ /ʃa/, /wʃa/ /ʃa/, /wʃa/ /ʃa/. Thus a phrase like /raj-tärjųj-rųj tsai-mutšįj-rųj is pronounced /raʃərjųj-shai-musirjə/. Whenever a morpheme ending in /ʃ/ immediately precedes a morpheme beginning with any consonant other than an alveolar or palatal obstruent, there is metathesis of /ʃ/ and the consonant. Morphemes do not end in consonants other than /ʃ/ (Doris Payne 1985b: 17). As in the Jivaro languages, nasal consonants are pronounced as prenasalized plosives,
I

that is with an oral release, in the environment of oral vowels. There are two contrastive tone levels; some syllables have inherent high tone; others inherent low tone; others have no inherent tone (1985b: 17–18).

The phonological systems of the Witoto languages vary considerably. Bora and Muinane have two contrastive levels of lexical tone. Tone is also very important in the morphosyntax of Bora, as shown in (3). Note, for example, that high tone occurs on the penultimate vowel in the citation form of ‘chief’, low tone when it is subject, and high tone on both the penultimate and antepenultimate when it is the possessor. Similarly the nominal form of mahtfo has high tone on the first syllable but the verbal forms do not, except in the future tense.

(3) (Thiesen 1996: 15)
áβéhú♣e ‘chief’
áβéhú♣e mahtfo ‘the chief eats’
áβéhú♣e mahtfo ‘the chief ate some time ago’
áβéhú♣e mahtfo ‘the chief is going to eat’
áβéhú♣e mahtfo ‘the chief’s food’

Each syllable has either a high or low tone and only one mora; the tones are the basis for sending messages a distance of several kilometres using a pair of hardwood signal drums about five feet each in length (Thiesen 1969). Muinane also has contrastive tone while Ocaina has contrastive pitch accent (Aschmann 1993: 3). Witoto, except for Marui, has contrastive accent (stress).

The phonemes of Bora are listed in tables 12.8 and 12.9.

All of the consonants except /r/ and /k/ have palatalized counterparts; /s/ and /s/ are the palatalized counterparts of /s/ and /s/. The phoneme /s/ functions morphophonemically as the palatalized counterpart of /s/ following /l/. Bora is unusual in that it has three close or near-close unrounded vowels.

---

12 Small language families and isolates in Peru

Table 12.8 Bora consonants

<table>
<thead>
<tr>
<th> </th>
<th>bilabial</th>
<th>alveolar</th>
<th>post-alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>labio-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>k[p, k*]</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>ts</td>
<td>tf</td>
<td>ts</td>
<td>tf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>ß</td>
<td>ß</td>
<td>x [s, c]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td>r</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>j</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ocaina also has palatalized consonants but not a full series. It is unique in that fortis and lenis nasals contrast. The fortis feature may be actualized either as length or as tense articulation or as both (Agnew and Pike 1957: 24). The lenis nasals are pronounced without complete closure (Leach 1969: 164). Voiced and voiceless plosives contrast. There are twenty-six consonants, five oral vowels, five nasal vowels and contrastive pitch accent.

Witoto is unique in that there is no voiceless bilabial plosive; Nipode does have [p] as an allophone of the voiceless bilabial fricative; like Bora, there is no [s]. Marui Witoto has a dental fricative /θ/. Nipode has voiceless implosives /ɓ/ and /ɗ/ word-medially, and Muinane has geminate plosives word-medially. Sequences of three vowels are common in Witoto.

The consonant and vowel inventories of Arbela from the Zaparo family are shown in tables 12.10 and 12.11 (adapted from F. Rich 1963).

---

Table 12.9 Bora vowels

<table>
<thead>
<tr>
<th> </th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>close</td>
<td>i [i ]</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>near-close</td>
<td>i [i ]</td>
<td>( \partial )</td>
<td>( \partial )</td>
</tr>
<tr>
<td>close-mid</td>
<td>o [o ]</td>
<td>( \partial )</td>
<td>( \partial )</td>
</tr>
<tr>
<td>open-mid</td>
<td>e [e , e ]</td>
<td>( \partial )</td>
<td>( \partial )</td>
</tr>
<tr>
<td>open</td>
<td>a [a , a ]</td>
<td>( \partial )</td>
<td>( \partial )</td>
</tr>
</tbody>
</table>

---

Table 12.10 Arbela consonants

<table>
<thead>
<tr>
<th> </th>
<th>bilabial</th>
<th>alveolar</th>
<th>post-alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>labio-velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>p</td>
<td>t [t , t]</td>
<td>k [k , k , k]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n [n , n , n]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>s</td>
<td>f [f , f]</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td>r [r , r , r , r]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td>j [j , j]</td>
<td>w [w , w]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Table 12.11 Arbela vowels

<table>
<thead>
<tr>
<th> </th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>close</td>
<td>i</td>
<td>i</td>
<td>u [u , u]</td>
</tr>
<tr>
<td>close-mid</td>
<td>o [o ]</td>
<td>a [a , a]</td>
<td></td>
</tr>
<tr>
<td>open</td>
<td>a [a , a]</td>
<td>a [a , a]</td>
<td></td>
</tr>
</tbody>
</table>
Zaparo differs from Arabela in having alveolar and postalveolar affricates and a glottal plosive. Iquito and Zaparo each has only four vowels: close front and central, open central, and one back vowel varying between close and mid. In all three languages long and short vowels contrast. In Arabela, at least, there is palatalization and labialization of consonants (written as Ci and Cu in the examples) in syllables following /i/ and /u/, respectively. In Arabela /h/ has a nasal quality and vowels are nasalized following /h/ and nasal consonants (F. Rich 1963: 197). With the exception of a phrase-final fortis/h/ and [ʔ] for emphasis in Arabela, syllables are open. A two-tone system is reported for Iquito (Eastman and Eastman 1963: 146).3

The isolates present several phonological features which differ from areal patterns. Ticuna has a very complex tonal system as well as laryngealized vowels (L. Anderson 1959a, b, and Montes Rodríguez 1995).6 Candoshi has a voiceless retroflexed post-alveolar affricate, and the maximum syllable is CCCVC. Taushiro is the only Peruvian language in which there are no bilabial or labio-dental consonants.

3 MORPHOLOGY

Word classes in most of the languages include open classes of nouns, verbs and adjectives. Closed classes usually include adverbs as well as pronouns and conjunctions. In most of the languages, adjectives are more likely to occur as predicate adjectives than in the noun phrase but the adjective class as such is not necessarily closed.

All of the languages of the five families are agglutinative, although strings of affixes tend to be shorter in the Jívaro languages than in the others. Suffixes predominate but there are causative prefixes in Jívaro, Cahuapana and Harakmbet. The Cahuapana languages also have about two dozen stem-forming verbal prefixes. In Zaparo, Bora and Yagua, person markers are prefixed to possessed nouns. In Yagua, Iquito and Harakmbet, and sometimes in Arabela and Zaparo, subject person markers are prefixes also. Cahuapana, Jívaro and Harakmbet verbs are suffixed to both subject/agent and object. The cross-referencing morphemes are prefixes in Harakmbet; declarative, dubitative and imperative sets differ in some subject/agent–object forms.

The Witoto languages and Yagua have dual, as well as singular and plural, numbers (a dual number is unusual in Amazonia). Cahuapana languages and Harakmbet have dual number in the first person only.

3.1 Nominal morphology

3.1.1 Classifiers

In Yagua, classifiers are infixed into the number 'one' and suffixed to other numbers; classifiers occur with nouns, demonstratives, adjectives and verbs. They are especially important in Bora where more than 350 have been identified. In Witoto and Peba-Yagua, the only families with gender distinction, masculine or feminine gender suffixes function as classifiers in words referring to animate nouns. In Bora classifiers occur in pronouns also. Examples from Bora with -ʔáami 'leaf' follow:

(4) (Thiesen 1996: 106–7)
\[\begin{align*}
\k'\text{aahák}^{\text{tu}} & \ 'knowledge' \ \\
\text{mit}^{\text{ba}} & \ 'big' \ \\
\text{ts}^{\text{ba}} & \ 'one' \ \\
\text{ts}^{\text{ba}} & \ 'another' \ \\
\text{a-} & \ 'conjunctiion, \ áʔáami \ 'that leaf, etc., mentioned in preceding sentence' \ \\
\text{p}^{\text{ba}i\tshí\text{t}} & \ 'important' \ \\
\end{align*}\]

In the Witoto family and Yagua, numerals agree with their head noun as in (5); in Bora and Yagua, demonstratives do also.7 The modifier in descriptive noun phrases may optionally agree, as in (6).

(5) (Yagua: Doris Payne 1985b: 114)
\[\begin{align*}
tá-nu-kii & \ nínuy \ \\
tá-cl.\text{thick. pole}- & \ \text{one pole} \ \\
\end{align*}\]

(6) (Yagua: Doris Payne 1985b: 117)
\[\begin{align*}
\text{rá-bii} & \ \text{rúna}(-bii) \ \\
\text{its-cl. flower} & \ \text{red-cl. flower} \ \\
\text{'its red flower'} \end{align*}\]

3 Further analysis might show that Iquito has a pitch-accent system.

4 Anderson (1959a, b) and Soares (1995) describe Ticuna as having five tones as well as glides. In a recent auto-segmental analysis, Montes Rodríguez (1995) agrees with Anderson's phonetic transcription but considers that high and low tones comprise the basic opposition and that a third mid tone is relatively weak and might best be considered as an underspecified element.

5 The data available to me are insufficient to determine if this is true of all Witoto languages or not.
In Cahuapana and Zaparo, one of the frequent noun derivational suffixes is an attributive ('owner of'), as in so?ja-wan (wife-ATTRIB) 'one who has a wife', ma?wan (things-ATTRIB) 'a rich person' (Chayahuita: Hart 1988: 261).

3.1.2 Case

Nomimative-accusative type case markers occur in all languages of the Cahuapana, Jivaro and Witoto families, and in Candoshi and Harakmbet. In Bora and the Jivaro languages, the nominative is zero. Witoto itself optionally marks both nominative and accusative (the object is obligatorily marked when the subject is third person). In Cahuapana both may be marked or both may be unmarked; the subject is usually marked only to avoid ambiguity or for emphasis and the object only for emphasis.

In Arabela, subject and object functions are usually marked by constituent order only but -ri may mark the subject when it precedes the verb. The -ri subject marker occurs primarily in transitive constructions and may be indicative that Zaparo languages are partially ergative. Another ergative-like feature in all of the Zaparo family is the fact that there are two sets of pronouns, the second of which could be considered absolutive since it is used for the object and for the subject of stative clauses. The use of one set or the other, however, depends primarily on constituent order; the first set is used for agents/subjects and also for objects when they precede the verb.

In Yagua, definite object enclitics attach to any word which follows the verb and directly precedes the object; or they attach to the verb when the object is not a fully specified noun phrase, as in (7d). Compare the agent and object references in (7).

(7a) (Thomas Payne 1983: 176–7)

*tsa-huuj-ma\_ Anita
[tsahuj mijad\_]
3sg.A-fall-PERFV Anita
'Anita fell.'

(7b) *tsa-puutji Pauro-nil Anita
3sg.A-carry Paul-3sg.O Anita
'Paul carries Anita.'

(7c) *tsa-puutji-nil Anita
3sg.A-carry-3sg.O Anita
'He carries Anita.'

(7d) *tsa-puutji-nil
3sg.A-carry-3sg.O
'He carries him/her.'

Thomas Payne glosses this example '... him/her/they' but he, Doris Payne and Paul Powilson all gloss -nil as '3sg'.

Thomas Payne (1983: 180) argues that this kind of system can lead to an ergative reanalysis and places Yagua typologically with those languages which utilize ergative construction types as the unmarked transitive construction in discourse, where the direct object is not being introduced into the discourse for the first time.

The cases in Murui Witoto are: -di 'nominative', -na 'accusative/oblique', -do 'instrumental', -ri 'dative', -ma 'allative', -mona 'ablative', -kon 'locative'.

In Jivaro, accusative and dative cases are not distinguished; the marker in Achuar is -ri-nil-an-un. Other Achuar cases are -hii 'instrumental, comitative', -ki-ak 'means', -numl-nam 'locative', -V 'genitive', -V 'vocative'.

In Zaparo and Yagua there are not only basic locative suffixes or postpositions but more than twenty-five postpositions indicating more exact location in time or space; many of them are compound, e.g. Arabela -koko 'on, above', -hi 'from', -koko-hi 'from above'; -koma 'below', -ra 'for', -koma-ra 'downwards'.

The Yagua applicative verbal suffix and the 'instrumental, comitative' case marker have the same shape and are no doubt etymologically related (Doris Payne 1985b: 178). This is also true for the Arabela suffix -tal-tla 'applicative or instrumental/comitative'.

3.2 Pronouns

Common to all five language families are personal, demonstrative, possessive and interrogative pronouns.

Singular, dual and plural are distinguished for all persons in the Witoto and Peba-Yagua families. In Cahuapana languages, dual is distinguished only for first person inclusive (and first dual inclusive also has an impersonal sense). First person inclusive and exclusive are distinguished except in Jivaro - where first person singular is used for exclusive and first person plural implies inclusive - and for the Witoto-Ocaima branch of Witoto. Demonstrative pronouns distinguish three degrees of distance and are often partially identical with third person pronouns. Singham can lead to an ergative as can be seen in the case of Murui Witoto in table 12.12 (from S A. Burch 1983, vol. II: 149–50). (Encliticized subject person markers are listed only if they differ from the free forms).8

Jivaro does not distinguish singular and plural in third person, and the demonstrative for second degree of deixis is the same as one of the third person pronouns, as shown in table 12.13 for Achuar (from Fast, Fast and Fast 1996: 31, 33).

Possessive pronouns in Jivaro and Cahuapana are formed by the addition of a case marker to the personal pronoun, e.g. ka 'I', kaki 'mine' (Chayahuita: H. Hart 1988: 262–3).

8 'Impersonal' is used in the sense of 'one (does)', i.e. an unspecified subject or agent.

9 The grave accent marks the final vowel of a diphthong in Witoto.
Table 12.12 Murui Witoto pronouns

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Dual</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>masc</td>
<td>kù</td>
<td>kâu</td>
</tr>
<tr>
<td></td>
<td>fem</td>
<td>kânu</td>
<td>kâu</td>
</tr>
<tr>
<td>2</td>
<td>masc</td>
<td>a</td>
<td>-ma</td>
</tr>
<tr>
<td></td>
<td>fem</td>
<td>-ma</td>
<td>kâu</td>
</tr>
<tr>
<td>3</td>
<td>masc</td>
<td>iâue</td>
<td>iâu</td>
</tr>
<tr>
<td></td>
<td>fem</td>
<td>iâue</td>
<td>iâu</td>
</tr>
<tr>
<td></td>
<td>neut</td>
<td>iâue</td>
<td>iâu</td>
</tr>
</tbody>
</table>

Table 12.13 Achuar pronouns

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>wi</td>
<td>ii</td>
</tr>
<tr>
<td>2</td>
<td>ami</td>
<td>âu</td>
</tr>
<tr>
<td>3</td>
<td>òamu</td>
<td>âu</td>
</tr>
</tbody>
</table>

3.3 Adjectives

In some of the languages adjectives are not clearly distinguished from nouns in that they may occur with most of the nominal affixes. However, when they function as nouns they often have a slightly different meaning, e.g. Achuar äpu ‘fat’ (as adjective) but ‘the chief’ (as noun); pisù ‘hard’ but pisù-ri hard-3sg ‘its hardness or firmness’ (Fast, Fast and Fast 1996: 35, 236).

3.3.1 Comparative and superlative grades

Comparative and superlative grades are not always distinguished. In Chayahuita the comparative/superlative is formed by reduplication of the adjective stem, e.g. noja ‘good’, noja noja ‘better/best’. The standard precedes the compared form, as in (9).

(9) (H. Hart 1988: 267)

ipi kiran wa?wi wa?w-ifin im-a-s??
‘That [animal] is smaller than the agouti.’

In Yagua the standard, which is a postpositional phrase, follows the comparative form, as in (10).

(10) (Doris Payne and Thomas Payne 1990: 288)

hârrij tsâmij Anita raj-janîhû
‘Anita is prettier than I.’

Comparison can also be made by juxtaposition of clauses, as in (11).
3.4 Verbal morphology

Verbal morphology in all of the families is considerably more complex than that of nouns and other word classes; so much so that in narrative text many clauses consist just of a verb.

In Jivaro and Cahuapana, suffixes referring to both the agent and the object occur in object-agent order in Jivaro and in the opposite order in Cahuapana. When there is no suffix agreeing with the object, it is third person, as in (12).

awir-in-ko ‘He hits me.’
awir-in-kin ‘He hits you.’
awir-in ‘He hits [him/her/them].’

In Iquito and Yagua a prefix agreeing with the subject/agent occurs, and in Witoto a suffix. In Bora the classifiers can occur as verbal suffixes agreeing with a third person subject/agent as in (13b). First and second person subjects are independent pronouns preceding the verb, as in (13a).

(13a) (Thiesen 1996: 58)
ó tsiiné ‘I ran.’
(13b) (Thiesen 1996: 108)
thèápo-ópe ‘he treats (the sick)’
tèápo-tse ‘she treats (the sick)’

Incorporation of the object is described for Cahuapana and Harakmbet. The object thus incorporated may be a noun stem or a classifier.

a?pi-naja-ti-r-awir to.light-eye-APPL-INDIC-1sg.A
‘I shone the light in (his) eyes.’

12 Small language families and isolates in Peru

Compound verbs are described for Bora (15), *Chayahuita* (16) and Harakmbet

(15) (Thiesen 1996: 60)
i2bhúdá-tuhskëénw-úpe / tuhskëénw-úpe
to.speak-to.begin-3sg.masc / to.begin-3sg.masc
i2bhúdá-ne
to.speak-action.in.general
‘He began to speak.’

(16) (H. Hart 1988: 271)
nito-ta?a-r-in
to.know-to.run-INDIC-3sg
‘He knows how to run.’

(17) (Tripp 1995: 205)
e?bëtë? + e?bëiër > e?bëjtëë?
INF-guard/possess INF-to.die
‘to guard a dead animal’

In Candoshi repeated action is expressed by repetition or reduplication of the verb root. In a series of individual actions the repeated root is followed by an inflected pro-verb as in (18); in continuous repetition the reduplicated verb root occurs in a finite verb, as in (19).

(18) (S. Tuggy 1982: 41)
Tpots kos kos kos ajira-g-ana
people to.arrive to.arrive to.arrive to.do.thus-CURR-PAST-3pl
‘A group of people arrived, then another group, then another group.’

(19) (S. Tuggy 1982: 41)
Tpots karowar-tar-tamta-e-ja
people to.descend-to.descend-HABITUAL-also-3-EMPH
‘A multitude of people descended also.’

*Chayahuita* distinguishes aspect but not tense in the verb. Jebero and languages of the other families have both tense and aspect suffixes. In Yagua, Jivaro and Cahuapana the unmarked tense is the present. Witoto itself distinguishes non-future and future tenses. Jivaro distinguishes five past tenses: immediate past, recent past, remote past, habitual past (over a long period of time) and reportive past. Jivaro also distinguishes immediate, definite and indefinite futures. Example (20) gives Achuar phrases with the irregular verb *tiiti* ‘to say’, and (21) gives further examples of contrasts in tense.
Mary Ruth Wise

(20) (Fast, Fast and Fast 1996: 75–7)

táwai ‘he says’
turútui ‘he says to me’
ti ‘he said (immediate)’
timiaji ‘he said (recent or remote)’
tinuji ‘he used to say’
táu ‘he said (reported)’
titatui ‘he will say (immediate/definite)’
títinuilai ‘he will say after some time’

(21) (Fast, Fast and Fast 1996: 45–6)

jááfá takákma-s-ma-mí ‘You worked yesterday (recent).’
amí-fa takákma-s-mia-mí ‘You worked some time ago (remote).’
jamáikia takákma-s-ta-hai ‘He has worked’
kafín wi takákma-s-tat-hai ‘He has made him work’

Aspects distinguished include perfective (or completive or terminative), durative (or progressive, incomplete) and iterative.

Directional suffixes occur in all of the language families except Jivaro. Iquito has a representative set: -wii ‘up, upriver’, -kuaa ‘down, downriver’, -kuwii ‘returning’, -sawii ‘arriving’, -maa ‘indefinite in time/direction’ (Eastman and Eastman 1963: 180). In Yagua, categories having to do with movement are ‘bounded’ (‘upon arrival’, ‘upon departure’, ‘en route’, ‘stationary’) and ‘unbounded’ (e.g. ‘while wandering around’).

In Jivaro, Cahuapana and Witoto desiderative affixes occur, especially if the subject of the complement and ‘to want’ are identical. The desiderative in Cahuapana is a prefix (ja- in Chayahuita).

(22) Muri'l Witoto (Burch and Wise 1968: 23)
kura-aka-d-e
see-want-THEMATIC-3sg
‘He wants to see.’

12 Small language families and isolates in Peru

All of the languages have morphological causatives. In Cahuapana and Jivaro languages and Harakmbet a causative prefix consisting of a single vowel is identified by David Payne (1990b: 78) as a wide-spread grammatical form in South American languages. These languages, as well as the others, also have causative suffixes. In most cases the causative affix can occur with both transitive and intransitive verb roots.

(23) Achuar (Fast, Fast and Fast 1996: 38–9)
wákí-tin vs i-wák-tin to.climb-INF CAUS-to.climb-INF ‘to climb’ ‘to lift’
takákma-s-mia-ji vs taká-mtik-sa-mia-ji to.work-PERFV-REC.PAST-3sg to.work-CAUS-PERFV-REC.PAST-3sg ‘he has worked’ ‘he has made him work’

In Witoto itself one of the causative suffixes is -ta. A valency-changing suffix with the consonant t is wide-spread in Amazonian languages (Wise 1993). In Zaparo, Peba-Yagua and Cahuapana (and possibly Jivaro although it has not been analysed as such), a suffix with t functions as an applicative11 so that normally oblique noun phrases are treated as direct objects. The applicative may also introduce another argument although it sometimes remains implicit. In Cahuapana the suffix -til-ta may verbalize, transitivize, change impersonal verbs to intransitive, detransitivize, and change transitive to ditransitive, as in (24) (-r ‘indicative’; -in ‘third singular agent or subject’; third singular object is zero).


ira ‘trail’ (noun) ira-ti-r-in ‘he/she walks’
amá-ti-in ‘he/she bathes’ ama-ti-r-in ‘he/she bathes him/her’
tafi-r-in ‘becomes night’ tafi-ti-r-in ‘it becomes night in the place where he/she is’ (lit. ‘it nights on him/her’) nati-ti-r-in ‘he/she obeys’ him/her
a?pa-r-in ‘he/she sends it’ a?pa-ti-r-in ‘he/she sends it to someone’

In Zaparo languages, the meanings of -tal-tia include: in a container, with contents, a group/plural, affection, perform the action carrying something (25), involuntary

accompaniment (i.e. being carried), to do with a goal/reason, the subject is sick/old/wounded (26), and reflexivizer, i.e. intransitivizer. An additional function, unique to Zaparo, is that it occurs in the main verb when the subject of the infinitival complement is different, as in (27).


kua morehaka tiurii-tia-ree-niia
1sg manioc stumble-APPLC-COMPL-1sg
'I stumbled while carrying my load of manioc.'

(26) (R. Rich 1999: 56)

hanija-ri nu-koko-hi kua fikiorta-aji
1sg-S trail-by-of 1sg hurt-CL
rofi-jo-ho-fa-rikio-wa-ni
drag-CONT-CONTACT-MULTIPLE-APPLC-IMPERFV-REPET-IR

With my wounded leg I paining me at each step, I dragged myself along the trail and returned [home].'

(27) Arabela (R. Rich 1999: 91)

hanija kia pani-tia-a kiu-nu-ni
1sg 2sg want-APPLC-CONT go-INF-IR
'I want you to go.'

Reflexive and reciprocal suffixes usually precede aspect and person suffixes, as in Achuar (28).

(28) (Fast, Fast and Fast 1996: 38)

tfarú-k-hai vs tfarú-ma-k-hai
cut-PERFV-1sg cut-REFL-PERFV-1sg
'I cut it.' 'I cut myself.'

In Bora there is one valency-reducing derivation which can be interpreted as reflexive or passive according to the context.

(29) (Thiesen 1996: 59)

ő kátatání-mu-mei
1sg cut-REFL
'I cut myself / I was cut.'

A suffix -j is reported for Yagua, which forms an intransitive stem from a transitive root with O becoming S.12

12 Doris Payne (1985b) uses the term 'anti-causal' for this suffix.

Evidentials occur in Bora and Zaparo. In Arabela there appear to be only reportative evidentials: -na is suffixed to the subject if it precedes the verb, as in (31); if there is no subject preceding, -kinia is suffixed to the verb, as in (32).

(31) (R. Rich 1999: 80)

kua ke-ja-na kua nikitio-ja kuno pueja-no-ni
1sg father-pl-REP 1sg give-CONT that person-sg-IR
'They say my parents are giving me to that man.'


kuso-ja-kinia
sick-CONT-REP
'They say he/she is sick.'

In Bora the reportative -tha indicates that the speaker was not a witness of that which he reports; the reportative -bá indicates that the speaker is reporting what he heard. The two may occur together, as in (33).

(33) (Thiesen 1996: 97)

Hotsée-~á-?há-ph e uymi~á kátatáí-76ó-ha-tw
Joseph-REP-unseen-PAST escaped dark-room-house-from
'Joseph escaped from jail a while back (the one who told me was not a witness).'

4 SYNTAX

4.1 Constituent order

The preferred constituent order is AOV, SV in the Jivaro, Witoto, Cahuapana families and in Arabela and Harakmbet. The preferred order in Iquito, Zaparo and Ticuna is AVO, SV; in Yagua and Taushiro it is VAO, VS. Urarina is the only Peruvian language in which it is OAV, SV. All of the languages have postpositions.

Demonstrative, quantitative and descriptive adjectives generally precede their
nouns, as in (34a); descriptive adjectives, however, can follow the noun, as in (34b). In compound nouns the modifier is usually first, as in (34c) and (35).

(34a)  
*Arabela* (R. Rich 1999: 94)  
keraatia pueretuuuka sapitaaaha  
'many large fish'  

(34b)  
kua mueruu kiari-nii  
1sg machete new-cl  
'my new machete'  

(34c)  
kahi-rikiako  
'the handle'  

(35a)  
Clzayallüu (H. Hart 1988: 261)  
pimo-Japon  
'fragrant-soap'  

(35b)  
finiti-niti  
to.be.thin/skinny-nose  
'narrow nose'  

In *Arabela*, descriptive adjectives occur much more frequently as the predicate adjective of equative clauses than in a noun phrase.

(37)  
(H. Hart 1988: 292)  
to?tiratin tjantjopí majowi?  
claws long very  
'It's claws (of the sloth) are very long.'  

Negative affixes usually occur in the verb, as in Murui *Witoto hikanxe-je-de* (ask-NEG-THETMATIC-3sg) 'he did not ask' (B. Burtch and Wise 1968: 23). Negative interjections are reported for all of the languages, e.g. Achuar *atsátsá* 'no'.

In Yagua there is no specifically passive construction; however, predicate nominal constructions often convey a passive sense, as in (40).

In *Witoto* itself, as in (41), and *Zaparo* languages, as in (42), there are specific passive constructions; the agent may be explicit in *Witoto*.

Interrogative words in content questions are usually the first element or cliticized to the first element in the sentence, as in *Zaparo* (38); in that language, however, interrogative particles in polar questions are always cliticized to the subject, as in (39).

(38)  
(Peeke et al. 1991: 14)  
tia ájtJonkí aí  
what name be.INTERROG 2sg  
'What is your name?'

(39)  
(Peeke et al. 1991: 13)  
maha-ká tja-ti äwiro-ka atsá-ka?  
raw-sg 2sg-INTERROG beetle-sg eat-CONT-ANTICIPATORY  
'Are you eating a raw beetle?'  

In *Witoto* itself, as in (41), and *Zaparo* languages, as in (42), there are specific passive constructions; the agent may be explicit in *Witoto*.

In Murui *Witoto* (B. Burtch and Wise 1968: 25)  
bie enurue  
this earth you to.make-PASSV  
'The earth was made by you.'  

(41a)  
Murai *Witoto* (B. Burtch and Wise 1968: 25)  
bie enurue  
this earth you to.make-PASSV  
'The earth was made by you.'  

(41b)  
hai  
phi-ka-mo  
mareña ha bie-mo  
already to.prove-PASSV when good then there-from  
good-FUT-INDIC-3sg  
'Now when it has been proved good, it will be good from there.'  

Most kinship terms and body parts are inalienably possessed; however, the possessive construction does not differ, whether the nouns are inalienable or alienable.

(42)  
paa tari huura-ha-fi-ja-ree-ni  
1.incl now disappear-CAUS-PASSV-CONT-COMPL-1R  
'Ve (our population) are being decimated.'
4.2 Relativization

Relative clauses generally follow their head nouns if there is one. In Chayahuita the only mark of relativization is -so? suffixed to the verb, e.g. iti awir-an-so? (agouti strike-2sg-REL) ‘the agouti which you struck/hit’ (H. Hart 1988: 264).

In Yagua, relative clauses – enclosed in square brackets below – consistently follow their head nouns; non-pronominal relativizers are formed with the demonstrative -hij plus the neutral classifier -ra and the enclitic -tij, as in (43a), or a relative pronoun may be formed by adding the enclitic -tij to a pronoun such as ríj, as in (43b). The verbs in both clauses are finite.

(43a) (Doris L. Payne 1985b: 69)

ramjitiwjerja himjitfara [hirjátiij] tsa-tárrjúj
raj-mutiwjej-ărá hij-ra-tij
1sg-cook-INAN food DEM-CL:NEUT-REL 3sg-buy
Tomásta-ra
Tom-INAN[RESUMPTIVE]
‘I cooked the food that Tom bought.’

(43b) (Doris L. Payne 1985b: 71)
néé tsámír-ja [ritimjúj tūwāätʃu tsimu]
ríj-tij-múj tsa-imu
NEG good-CL:NEUT 3pl-REL-NEG listen 3sg-LOC
‘Those who don’t listen to him/her are not good.’

The common argument may have any function in the main clause and in the relative clause.

In Jivaro languages there are no relative pronouns but third person pronouns may have this function. The more common construction is simply a nominalized clause, as shown in the Aguaruna examples in (44).

(44a) Aguaruna (Larson 1978: 296, 299)
takatsusi [suhumainu-k]
we.do.not.carry.in.hand what.might.be sold-TOPIC
‘We do not have anything to sell.’

(44b) Aguaruna (Larson 1978: 252, 255)
máktʃik aints [matʃiŋki-n taŋkumau]
one.person monkey-O one.who.tamed
’a person who tamed a monkey’

In Bora an independent clause with the subject marked by a pronominal suffix may function as a subject, as in (45b), or object, as in (45c); or if there is a nominal subject, the inanimate marker -nel-REL is suffixed to the otherwise independent verb in the relative clause, as in (45d).

(45a) (Thiesen 1996: 91–2)
mitleased kš'ak'ime-ipl-e
much work-3sg
‘He works hard.’

(45b) [mitleased kš'ak'ime-ipl-e tsháhúk̓ oó]
much work-3sg -come
‘The one who works hard is coming.’

(45c) [mitleased kš'ak'ime-ipl-e-k'eh] ó aʔtó
much work-3sg-ACC 1sg pay
‘I paid the one who works a lot.’

(45d) ó ahtšumú [Jódá kš'ak'ime-ml-e]
1sg saw John work-INAN
‘I saw John working.’

In Iquito, relative clauses are introduced by the demonstrative pronoun iíná ‘this, that’.

(46) (Eastman and Eastman 1963: 167)
káá ki-nikii kiaa-mutúúra [iíná kiaa-miíjáákura]
NEG 1sg-see 2sg-motor that 2sg-used.to.have
‘I don’t see the motor which you used to have.’

4.3 Subordination and coreference/switch-reference

Other nominalized clauses include complements and purpose clauses. When the subject of the verbal complement is the same as that of the main clause, the complement is normally nominalized with an infinitive suffix as in Arabela (47). (Compare (27) for use of the applicative in the main verb if the subjects are different.)

(47) (R. Rich 1999: 91)
hanija kia-ta kia-ну pani-ja-ni
1sg 2sg-comit go-INF want-cont-1R
‘I want to go with you.’
In Yagua if the subjects are different a finite verb is used in the complement with a non-coreferential person-marking clitic, as in (48).

(48) (Doris L. Payne and T. E. Payne 1990: 337)

1: hij-ñá (tsúñúj)
2: hij-wááa tsa-hünúúj Tomássa
3: 2sg-want 3sg-look Tom

"Do you want Tom to look/write?"

In Yagua, infinitival nominalizations have adverbial functions also. Usually they are suffixed with a postposition and person-marking clitics (equivalent to possessor reference), as in (49–50). If subject reference is omitted, it is understood that the subject of the infinitive is the same as that of the main verb.

(49) (Doris L. Payne and T. E. Payne 1990: 337)

1: tsatsíímjaa himjuut jääñññú-nii
2: tsa-tsiij-maa hij-himúùta-ha-nú-nii
3: 3sg-run-PERFV COREF-help-INF-ADLATIVE-3sg

"He has run to help him." (lit. 'he, has run towards him/)

(50) (Doris L. Payne and T. E. Payne 1990: 337)

1: [já] [sáá] [simbjeenúññájura]
2: hij-á tsa-himúùta-hanú-hú-nii
3: 2sg-want 3sg-look/INAN

"Give it to him to eat." (lit. 'give it to him, towards him, eating')

The Yagua coreference clitics indicate 'coreference with a previously mentioned subject, possessor or object of postposition within the same clause. . . . Part of the function of the set II [suffixal] coreference clitic is to indicate reflexivity and reciprocility' (Doris L. Payne and T. E. Payne 1990: 361).

Cahuapana and Jivaro, on the other hand, have subordinating verbal suffixes which indicate coreference or lack of it, i.e. switch-reference, relative to the agent or subject of the adverbial subordinate clause and that of the matrix clause.

In Chayahuita, switch-reference suffixes occur only in subordinate verbs with third person agents or subjects, as shown in table 12.14 (from H. Hart 1988: 489) with the sequential action, subordinate forms of the intransitive verb sakatirín 'to work'.

Switch-reference systems are highly developed in Jivaro languages. Long chains of subordinate clauses occur. Larson (p.c.) reports a complete Aguaruna narrative consisting of 136 subordinate clauses, including embedded dialogue, marked for

same or different agent or subject; these are followed by one independent clause. A person marker occurs in the verb, and the subordinating suffixes also mark temporal and or logical relations. The subordinate clauses are marked by -k/aj 'topic marker' if they are part of the main event line.

BIBLIOGRAPHY


Pucallpa: Instituto Lingüístico de Verano.


Trudell, B. 1993. Beyond the bilingual classroom: literacy acquisition among Peruvian Amazon communities, SIL and the University of Texas at Arlington Publications in Linguistics, 117. Dallas: SIL and University of Texas at Arlington.


Walton, James. 1977. 'Participant reference and introducers in Muinane clause and para-
Other small families and isolates

ALEXANDRA Y. AIKHENVALD AND R. M. W. DIXON

The other small families and isolates are grouped in this chapter, according to the country in which most of their speakers are located; some spill over into a nearby nation. Table 13.1 summarizes the languages, estimated number of speakers and approximate locations. These locations are given on map 12. (A number of other languages, now extinct, on which little or no information is available, are not listed here.)

1 BRAZIL

There are about 170 indigenous languages in Brazil. They belong to about 20 language families and are spoken by approximately 155,000 people (A. D. Rodrigues 1986). Languages from the larger families are discussed elsewhere in this volume—Carib in chapter 2, Arawak in chapter 3, Tupi in chapter 4, Tupi-Guarani in chapter 5, Macro-Jê in chapter 6, Tucano in chapter 7, Pano in chapter 8, Makú in chapter 9, Nambiquara in chapter 10 and Arawá in chapter 11. Languages of the Chapacura family, the Yanomami dialect cluster, and isolates Trumai, Mura-Pirahã, Jabutí, Maku and Aikanã are spoken exclusively in Brazil. Their typological properties are briefly discussed here.

1.1 Yanomami

The Yanomami (also known as Yanoama, Yanomami-Waicá, Xirianá or Guharibo) are one of the largest groups of non-acculturated tropical forest Indians in northern Brazil and in southern Venezuela (Migliazza 1972; Tovar and De Tovar 1984: 163–4; Lizot 1988: 489). Yanomami constitutes a dialect continuum

Special thanks go to Nádia Pires, for providing information on Jabuti, Henri Ramirez for data on Yanomami, Hein van der Voort for data on Koasá and Ione Vasconcelos, for information on Aikanã. §13.1.2, on Trumai, is a condensation, by the authors, of materials written by Raquel Guirardello. We are grateful to Paul Frank, Francisco Queixaílos, Terry Malone, Janet Barnes and Tim Curnow for commenting on a previous draft of this chapter.
which could be regarded as making up a single language. There are four main dialect groups as set out in table 13.1.

Local groups or villages identify themselves with names ending in -theri ‘dwellers of, village of’. This morpheme is suffixed to names of mountains, or rivers, or places named for particular vegetation or animals, e.g. jama-ti-i-theri (tapir-place-river-people) ‘Xamatauteri’ (a group of Yanomami: Ramirez 1994), aiva-tha-theri (aiwa-mountain-place-people) ‘a group of Yanam (on the Quaimi River, upper Uraricá)’ (Migliazza 1972: 32–3).

Table 13.1 Minor languages and isolates (Languages for which no linguistic information is available are marked with *)

<table>
<thead>
<tr>
<th>BRAZIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Yanomami dialect cluster</td>
</tr>
<tr>
<td>1a Yanomami (or Yanomai, Yanomae), c. 15,000. There is a northern subdialect in the Padamo river basin; a central one on the Ocamo, Mamaviche and upper Orinoco rivers; and a southern one south of the Orinoco up to the headwaters of the Marauia and Cauahurí.</td>
</tr>
<tr>
<td>1b Yanomami, c. 3,000. There is a southern subdialect around the upper Uraricuera and lower Parima rivers; a central one on the upper Parima river and the headwaters of the Orinoco; a southern one in the Catrimuni and Demeni rivers basins, and possibly a fourth on the Upper Ajarani river.</td>
</tr>
<tr>
<td>1c Yanam (or Ninam), c. 360. Its southern subdialect is spoken on the Mucuají river, and the northern variety is located on the upper Uraricá and Paragua rivers.</td>
</tr>
<tr>
<td>1d Sanuma (or *Tsanima), c. 500 in Brazil and c. 1,500 in Venezuela. One subdialect is on the Caura river up to the Brazilian border; another is on the Erevato–Venturi rivers; and the third is in the Auaris river basin.</td>
</tr>
<tr>
<td>2 Trumai, c. 50. Spoken in the Xingu reserve, on the Curiáse river, a tribuary of the Xingu river.</td>
</tr>
<tr>
<td>3 Mura-Pirahã, c. 250. Maeli river (off the Marmelos river, off the Madeira river), Amazonas.</td>
</tr>
<tr>
<td>4 Jabuti (Quipiu, or Jeromixi), c. 60 people in the indigenous area of Guaporé, Rondônia (the nearby language Arikapã, with just a few speakers, may be closely genetically related).</td>
</tr>
<tr>
<td>5 Chapacura family</td>
</tr>
<tr>
<td>5a Wari (or Oro Wari, Pokaa Nova, Pacaca Nova), c. 1,800. Along the tributaries of the Pacaca Novos river and other tributaries of the Mamoré river, Rondônia (dialects include ‘Oro Nao’ and ‘Oro Eu’).</td>
</tr>
<tr>
<td>5b Oro Win, c. 40. Around the headwaters of the Pacaca Novos river, Rondônia.</td>
</tr>
<tr>
<td>5c Uruá, c. 150. On the Uruá river, Rondônia.</td>
</tr>
<tr>
<td>5d Torã, c. 250. On the Marmelos and Paricá rivers, and on the Fisto Cabeça d’anta, Amazonas.</td>
</tr>
<tr>
<td>5e More, spoken by a few people in Rondônia and in Eastern Bolivia. A few more extinct languages belonging to this family are mentioned by Loukotka (1968: 160–2).</td>
</tr>
<tr>
<td>6 Maku, 3 speakers in 1964. Around the Uraricuera river, Roraima.</td>
</tr>
<tr>
<td>7 Aikéná (or Tobarão, Hauri, Massaká, Kasupá and Mundé), c. 120. To the east of the Upper Pimenta Bueno, Rondônia.</td>
</tr>
<tr>
<td>8 Koaia (Kwa3a, Arara), spoken by a few people who live together with the Aikéná.</td>
</tr>
<tr>
<td>9 *Irantxe (Mynky, Munkí), c. 190. On the Upper Jurua river, Mato Grosso.</td>
</tr>
<tr>
<td>10 *Sape (Caliama), 5 speakers in 1964. On the Upper Paraguá and its tributary Caroni; South Venezuela/Brazil border (Migliazza 1978a).</td>
</tr>
<tr>
<td>12 *Katukina, a family which includes Katuamari, c. 600; Katukina do Biá, c. 250; Txunhá-djapá, c. 30; and Katawixi, c. 10. Around the Upper Juruá, Jutai and Javari, Amazonas (A.D. Rodrigues 1986).</td>
</tr>
</tbody>
</table>
**Table 13.1 (cont.)**

<table>
<thead>
<tr>
<th>BOLIVIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 Tacana family</td>
</tr>
<tr>
<td>13a Arauca, c. 80. Around the headwaters of Mansapari river.</td>
</tr>
<tr>
<td>13b Cavineña (or Kavineño), c. 2,000. South-east of Riberalta, along the Beni river.</td>
</tr>
<tr>
<td>13c Tacana, c. 3,500. On the Beni and Madre de Dios rivers (some of the Tacana dialects were called Tiintingu - Girard 1971: 19).</td>
</tr>
<tr>
<td>13d Ese Eja (or EseEja, Chuna, Tiintingu, Huayray, Tamobopata-Guarayo; different names may correspond to different dialects - Girard 1971: 19-20), c. 600, on the Beni and Madre de Dios rivers. (and c. 250-400 in the regions of Tamobopata and Heath rivers around Maldonado in Peru).</td>
</tr>
<tr>
<td>13e Reyesano (San Borjano; also known as Sapibocona, Maropa; possibly Warísa - Girard 1971: 18-20; these names may correspond to different dialects), probably extinct. Around 1,000 members of the ethnic group reside in Beni Department, around San Borja.</td>
</tr>
<tr>
<td>Other Tacana languages, possibly extinct, include (Girard 1971: 17-20; Loukotka 1974):</td>
</tr>
<tr>
<td>17a Arasa, formerly spoken on the Marcopata and Arasa rivers;</td>
</tr>
<tr>
<td>17b Mabenaro, formerly spoken on the Matupiri river.</td>
</tr>
<tr>
<td>14 Cayuwana (Cayubaba, Cayuwa), nearly extinct. Formerly spoken west of Mamoré river and north of Santa Ana (there are reported to be 900 people in the ethnic group).</td>
</tr>
<tr>
<td>15 Ionama (Machoto, Sramo), c. 100 (all bilingual in Spanish). Ionama river.</td>
</tr>
<tr>
<td>16 Movima, c. 1,000. Around Santa Ana on the Yacuma river.</td>
</tr>
</tbody>
</table>

**COLOMBIA**

<table>
<thead>
<tr>
<th>Guahibo family</th>
</tr>
</thead>
<tbody>
<tr>
<td>17a Guahibo (or Sikuani), c. 20,000. Eastern Meta and Vichada and across the border in Venezuela.</td>
</tr>
<tr>
<td>17b Guayabero, c. 800. On Guaviare river.</td>
</tr>
<tr>
<td>17c Cuiba (or Cuiba-Wámonae), c. 2,000. Casanare, Arauca, Vichada and across the border in Venezuela.</td>
</tr>
<tr>
<td>17d Macuquín (or Jinju), c. 180. In the Department of Arauca, bordering Venezuela.</td>
</tr>
<tr>
<td>17e Payero, c. 150-60. Arauca river, between Arauquita and Arauca.</td>
</tr>
<tr>
<td>18 Sáliba-Piaroa putative family</td>
</tr>
<tr>
<td>18a *Piaroa, c. 80 in Colombia and c. 12,000 in Venezuela. South bank of the Orinoco river, inland from the Paguaza river to Manipiari, Amazonas.</td>
</tr>
<tr>
<td>18b Sáliba, c. 2,000. The Eastern Plains along the Meta and Casanare rivers and across the border in Venezuela.</td>
</tr>
<tr>
<td>19 Andoké, c. 75-100. The tributaries of the Caquetá.</td>
</tr>
<tr>
<td>20 *Pumane, c. 2,000. Intira river (and c. 240 across the border in Venezuela).</td>
</tr>
</tbody>
</table>

**VENEZUELA**

| Yaruco (or Liaruru, Pumé, Yuapir), c. 2,000-3,000. On the Orinoco, Meta and Apure rivers, in the states of Amazonas and Apure. |
| 22 Warao (or Guaraano, Guarao), c. 18,000. On the delta of the Orinoco. |
| 23 *Hoti (or Hodí, Yuwuna, Warowaru, Chicano), c. 300-500. On the Kaiwa river in Amazonas, and on tributaries of the Asita and Marupire rivers. |

The dialects are lexically very close, Sanuma being the most divergent. The average number of lexical cognates on a 200-word list is 70-80%; the dialects share 95% of their grammatical morphemes.

The Yanomami live in a mountainous jungle area of difficult access, away from main rivers. They live in small communities each of 20 to 350 people, spread over a total area of about 200,000 square kilometres.

Means of subsistence are hunting, fishing, gathering and slash-and-burn agriculture. The Yanomami are known as fierce, dangerous people and are feared by their neighbours (Chagnon 1992; Lizot 1988). They are reported to practise ritual cannibalism: when an important person dies, their body is burnt, then the bones are crushed in a pestle and mixed with banana mash. A feast is made to which people come from distant Yanomami villages. At this feast, men and children consume the mash and take hallucinogenic drugs. The Yanomami are constantly engaged in intervillage warfare; they are known to raid neighbouring villages, sometimes killing all the men, and capturing women and children.

The first contact with Europeans is reported to have been after 1800 (Migliazza 1985: 27). A few missions have been established in Yanomami-speaking communities, starting from 1950. Since 1970, their territories have been under constant threat of being invaded by non-indigenous settlers. In recent years the Yanomami on the Venezuelan-Brazilian border have suffered the effects of a massive gold rush (starting from 1987, see Gomez 1990:10). The number of Yanomami has fallen drastically during the past few decades.

The Yanomami are mostly monolingual, with a certain degree of polyglossia in different Yanomami dialects (Migliazza 1972). A ritual lingua franca, called 'wayamo', is spoken between people from different villages, mostly on festive occasions (1972: 47-58). About a quarter of Sanuma speakers are bilingual with Yekwana (or Dekwana), a Carib language. Nowadays, men tend to know some Portuguese or Spanish.

The information available on Yanomami languages varies in quality and in quantity. Migliazza (1972) provides an excellent description of the Yanomami dialect continuum, cast in a generative framework. Different varieties of Yanomami are described by Ramirez (1994) and Lizot (1996); there is a good grammar of Sanuma by Borgman (1990); and a short study of Yanam by Gomez (1990). (All the examples are from Yanomami unless indicated otherwise.)

1.1.1 Phonology

There are seven vowels - i, e, a, O, u plus high-central i and mid-central a. The consonant system is given in table 13.2.

In Yanam, u has fallen together with i. There is regressive vowel assimilation
Table 13.2 Yanomami consonants

<table>
<thead>
<tr>
<th></th>
<th>In all four groups:</th>
<th>Yanomami also has:</th>
<th>Yanonam also has:</th>
<th>Sanuma also has:</th>
<th>Yanam also has:</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>p, t, k</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>s, h</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>th</td>
<td></td>
<td></td>
<td>f</td>
<td>ts</td>
</tr>
<tr>
<td>nasal</td>
<td>m, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>w, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

across morpheme boundaries. Vowels are weakened in a non-accented word-final position and may be deleted.

Syllable structure is (C₁)(C₂)C₃V. The possibilities at C₁, C₂ are pr, kr, mr, he. Primary stress is assigned to the penultimate syllable of a word, and secondary stress to every alternate syllable before that.

Every morpheme is either intrinsically nasal, or intrinsically oral (Migliazza 1972: 157-9). A nasal root spreads its nasality onto an intrinsically oral suffix provided that it does not start with a stop or a fricative. There is no nasality spreading onto a clitic. Roots consist of two moras. Lexical items which contain three or more moras are interpreted as containing a two-mora root and one or two fossilized derivational affixes. The following morphemes are longer than two moras; their third syllable, which starts with a stop and thus 'blocks' nasalization from spreading, is treated as a suffix: dēkō 'termite', mōhōti 'idiot', mōhēko 'knee' (Ramirez 1994: 68-70).

1.1.2 Word structure and word classes

Yanomami is suffixing and predominantly head-marking with elements of dependent-marking. It is highly polysyntactic. Open classes are nouns and verbs; adjectival concepts are expressed with stative verbs. Pronouns and demonstratives are similar to nouns in that they take the same case markers. There are five demonstratives, the choice of which depends on distance from the speaker and from the bearer, and on the visibility of the object.

1.1.3 Nominal categories

There is a distinction between alienable and inalienable possession. Alienable possession is marked by simple juxtaposition of terms, e.g. Sanuma kamīsa sai-a (my house-sg) 'my house' (Borgman 1990: 128). Inalienably possessed body parts are obligatorily incorporated into the verb. They are positioned before the verbal root, after the cross-referencing markers. (1a) shows the incorporated noun ha 'head of'; (1b) is ungrammatical (Ramirez p.c.).

(1a) Jama ja-he-va-ma
    tapir 1sg-head-eat-PAST
    'I ate tapir's head.'

(1b) *Jama he ja va-ma
    If a noun functions as a predicate, the body part is encliticized to it, as in (2).

(2) Jama-he
    tapir-head
    'This is a tapir's head.'

Possessive relations can also be marked with independent pronouns or 1st and 2nd person possessive suffixes (1st -ja, 2nd -hp). 3rd person possessor is marked by a clitic -e, or by incorporating the possessed noun, with a suffix -pi, onto the verb.

Case markers are elided to the last word in an NP. They are -ni (ergative (with animate nouns); instrumental (with inanimate nouns)); -fo 'associative with'; and a general oblique case with the following forms: -ha 'inanimate, non-peripherical (near)', -ha 'animate, non-peripherical (near)', -ha-mi 'animate, peripherical (far)', -ha-ri 'animate, peripherical (far)'). This is used to express a variety of meanings (locative, directional, addressee, agent of causative, comparative – see Ramirez 1994: 193–200; and the comparison of case markers in Migliazza 1972: 331). In (3), -tha marks the beneficiary.

(3) vāro-pa-tha
    hiima ja-e-ki-pi-hip-o-ke-ma-he
    man-PL-BENEFICARY dog 1sg-OBL-PL-give-FOC-PAST-3PL.OBL
    'I gave the two dogs to the men.'

1.1.4 Classifiers

There is a rich system of verbal classifiers; some dialects are reported to have over 100. Classifiers categorize the S, as in (4), or the O, as in (5); they are obligatory and appear immediately before the verb root (Ramirez 1994: 127).

(4) kori una-yā-i
    crest CL:CREST-fly-DYNAMIC
    'The crest (bird) flew away.'

(5) kori-una-pa-ni
    mau u-kōa-he
    crest-CL:CREST-PL-ERG water CL:LIQUID-drink-DYNAMIC-3PL.ERG
    'The crests (birds) drink water.'
In (5) the verb includes the 'liquid' classifier, referring to the O argument. In (4) the verb includes a classifier referring to the S argument; this is a 'unique' classifier used just for 'crest' (a type of bird). In verbless clauses, classifiers are postposed to the predicate noun, as in:

(6) parava-u
river-CL:LIQUID
'This is a river.'

Some examples of classifiers are ko 'round'; si 'objects with large surface, palm trees'; ma 'hard objects'; ki 'snakes'; pa 'turtles'; u 'liquid', etc. Classifiers can co-occur with the plural marker -pa, but not with the singular marker -a. They are also used on nouns, e.g. kori-una-pa (crest-CL:CREST-PL) 'crests'. Most classifiers originate in body-part terms.

1.1.5 Grammatical relations

All Yanomami dialects have ergative case marking. In addition, the verb always cross-references A and O of a transitive, and S of an intransitive clause.

In (7), suffix -he cross-references an NP in the A function (marked with the ergative case), vâro-pa, while the enclitic -a cross-references the O, absolute fama 'tapir' (Ramirez p.c.).

(7) vâro-pa-niA fama0 a-nia-ma-he
man-PL-ERG tapir S/GABS-shoot.arrow-PAST-3pl.ERG
'The men shot arrows at the tapir.'

In (8), the absolute ('tapir') is unmarked, and is cross-referenced on the verb with absolute a.

(8) fama8 a-toku-ma
tapir S/GABS-run-PAST
'The tapir ran away.'

Cross-referencing markers indicate person and number (singular, dual or plural) of A, and just number for O or S. The distinction between inclusive and exclusive for 1st person pronouns has been lost in Yanam and Yanomam.

Oblique constituents (most often the beneficiary) can also get cross-referenced on a transitive verb. In (3) the oblique marker he, formally identical with the 3pl ergative marker, cross-references the addressee ('men') on the verb (further discussion is in Ramirez 1994: 315).

Yanam (Gomez 1990: 48ff., 129) is reported to have a split-ergative system depending on the semantics of NPs. In agreement with the Nominal Hierarchy, there is a nominative-accusative system for first and second person singular pronouns. The ergative marker is obligatory with third person free pronouns and with nouns.

1.1.6 Verb structure

Verbs may be intransitive, transitive, ambitransitive of type S = A or ambitransitive of type S = O. Verbal morphology is extremely complicated. There are more than twenty positional classes of proclitics (cf. Ramirez 1994: 100ff.) and over twenty positional classes of suffixes. Example (9) shows a complex verb, with four proclitics and eight suffixes.

(9) fama e-ki-pats-va-heri-pi-pra-me-re-ha-no-o-i
'making two huge tapirs rush away'

Proclitics are prepended to the verb but postposed to the noun in a verbless clause. They include pronominal cross-referencing, classifiers, number of O (dua dual, plural, collective), incorporated body parts, some evidentiality, modality, aspect, topic and focus markers, a diminutive and an augmentative.

The proclitic kâi marks a comitative type of applicative. The verb root hu- 'walk' is intransitive, while kâi-hu- 'walk with' is transitive, as in:

(10) (kamijo-ni) Hiterawa ja-kâi-hu-i
1sg-ERG Hiterawa 1sg-COMIT-walk-DYNAMIC
'I am walking-with Hiterawa.'

Verbal suffixes are used to mark number, causative, tense-aspect (perfective, dynamic, continuous, iterative, habitual, etc.), direction and negation. An intransitivizer -o is used to form reflexives, e.g. nia- 'shoot', nia-o 'shoot [arrow] at oneself'.

Causative -ma can be used on any verb (transitive or intransitive). If this suffix is added to a transitive verb, the causer takes the ergative case, the O remains O, and the underlying A is marked with the peripheral clitic (-iha). Example (11) shows a double causative (Ramirez 1994: 247).

(11) Hiterawa-iha ihiro ja-e-hiima-ara-me-ma-re-ma
Hiterawa-PERIPHERAL child 1sg-OBL-dog-run-CAUS-CAUS-TELIC-PAST
'I let Hiterawa make the dog of a child run.'

Yanomami has a four-term evidentiality system. 'Eyewitness' is shown by a suffix, 'deduced' and 'reported' by proclitics, and 'assumed' by a combination of
proclitic plus suffix. Sanuma (Borgman 1990: 165-73) has a three-term system, with 'eyewitness', 'verification' and 'supposition', while Yanam just has 'eyewitness' and 'non-eyewitness'. Different evidentiality markers are used in the three dialects, perhaps indicating a relatively late development of evidentials (also see Migliazza 1972: 268).

1.1.7 Noun incorporation and verb compounding
Any noun in S or O function can be incorporated into the verb to mark a topically continuous participant which is not in focus (Ramirez 1994: 385). Compare (12), without incorporation, and (13), where the direct object, 'axe', is incorporated. Incorporation does not affect the transitivity of the verb.

(12) (kamija-ni) sipara ja-puhi-i
   lsg-ERG axe lsg-want-DYNAMIC
   'I want an/the axe.'

(13) (kamija-ni) ja-sipara-puhi-i
   lsg-ERG lsg-axe-want-DYNAMIC
   'I want [it], the axe.'

Verb compounding is very productive. Compounded verbs consist of an independent verb root followed by a bound verb root in a modifying function, e.g. kô 'do again' in (14) (Ramirez 1994: 330-47). Bound verb roots cannot form a predicate by themselves.

(14) ja-pa-naka-i-kô-o
   lsg-PL-call-DYNAMIC-do.again-INTR
   'I called them again.'

1.1.8 Syntax
The unmarked constituent order is AOV, SV. There are no copula verbs. Yanomami has considerable freedom in the order of core and oblique nominal constituents within a clause. In Sanuma, constituent order displays ergative characteristics. Using X to indicate a peripheral constituent, the preferred order is AXOV, XSV, i.e. S and O pattern together, occurring between X and V.

Demonstratives, quantifiers, numerals and classifiers precede the head noun, while adjectival modifiers follow it.

Subordination is usually marked with verbal suffixes; some of these suffixes indicate switch-reference. Relative clauses follow the head. In Yanomami, Yanam and Sanuma a nominalizer can also mark the predicate of a relative clause or of a complement clause. Thus (Migliazza 1972: 86):

(15) wáro-n fama jíra-wei ware-ma
    man-ERG tapir kill-REL eat-COMPL
    'The man who killed the tapir ate.'

Sanuma also has a relative pronoun i. This dialect is reported to have internally headed relative clauses; the main clause contains a pronoun coreferential with the common argument. See (16), from Borgman (1990: 133). There are no restrictions on the functions of a common argument.

(16) [i:pá uku kôkô lo-le] i kôkô hu mai kite
    my son 3pl sit-PRES REL 3pl go NEG PUT
    'My sons who are sitting here [lit. 'they'] will not go.'

The type of marking of a relative clause depends on its predicate. If the predicate of a relative clause is an active verb, the relative pronoun can be omitted, and just -wi 'relativizer, nominalizer' is used. If the predicate of the relative verb is a stative verb, the predicate takes -i 'relativizer' as in:

(17) [i:pá hoose a pata lape epi-i]
    my younger.brother 3sg AUGMENTATIVE tall INTENSIVE-RELZR
    t a niha sa te toto-ki ke
    REL 3sg to 1sg 3sg give-FOC IMM.PAST
    'I gave it to [lit. 'him'] my younger brother who is tall.'

1.2 Trumai by Raquel Guirardello
The Trumai say that they came from the south-east into the Upper Xingu region (see chapter 15) within quite recent times. They provide an unusual example of (temporary) language revival. Trumai numbers were reported to be 43 in 1928, 25 in 1948, only 18 in 1952, then 21 in 1963 and over 100 (including many children) by the early 1990s. However, only about half (approximately 50) of the ethnic Trumai speak the language today. There has been a great deal of marriage with people from other Xingu tribes, allied to the fact that the other Xingu people consider Trumai 'too hard' a language to learn. Owing largely to contact with a nearby government post, most young people have switched to Portuguese. It is not likely that the language will continue to be spoken for more than another couple of generations.

The first vocabulary of Trumai was taken down by von den Steinen (1886). Monod-Beccuelin (for example, 1975, 1976) attempted some grammatical analysis. Guirardello (1992) is a workman-like preliminary statement of the phonology and morphology which will be refined in Guirardello (forthcoming).
negative, causation and emphasis. Negation is shown by anuk in copula and tak in non-copula clauses. There is a 3rd person enclitic -nl-e (referring to an argument in S or O function) which is attached to the right of the verb phrase; interestingly, this cannot co-occur with a negative marker.

For alienable possession the possessor bears the genitive suffix -k(a)te and precedes the possessed, e.g. Kumaru-kte tahu 'Kumaru's knife'. Inalienable possession can be shown just by apposition, e.g. dinoxo kuf 'the girl's hair', ha kuf 'my hair'. When a third person (anaphoric) possessor pronoun is employed, inalienably possessed nouns subdivide into two classes. There is a prefix t(s) - on kin terms, e.g. tsipine 'his/her cousin'; and there is an enclitic -aki (or -ea preceding a postposition) on body-part nouns and also names for clothing, home, shadow, name and pet, e.g. kuf-ake 'his/her hair', kuf-ea letsi 'with his/her hair'.

The basic clausal constituent order is SV, AOV, with a dative NP following the verb. If the S or O NP does not immediately precede the verb, a particle ke must be included after the verb. In contrast, an A or dative NP can be moved around without any post-verbal marking. Within an NP, quantifiers and possessors precede and adjectives follow the head. Within coordinate constructions, coreferential deletion can operate on an S = O or S = A basis, suggesting a pragmatic (rather than a syntactic) pivot.

The causative of a transitive clause has a most unusual structure, in which both the causer and the A of the underlying verb take ergative marking (the ergative enclitic is -ts after 1sg and -ela/k elsewhere).

13 Other small families and isolates

Trumai has a dependent-marking language. An NP in S and O function takes no marking and there are postpositions for ergative (A function), dative, locative, ablative, instrumental and comitative (the ergative, dative and locative markers are enclitics). Verbs fall into four classes: plain intransitive (obligatory S argument), plain transitive (A and O), extended intransitive (S and dative) and extended transitive (A, O and dative). Verbs in the extended intransitive class include 'eat', 'drink', 'see', 'like' and 'talk with'. There are two verbs 'kill' - fa which is extended intransitive and disi which is transitive.

Trumai has little morphology. There are no TAM suffixes to the verb, reference to time being expressed through adverbs. There are post-verbal particles marking

### Table 13.3 Trumai consonants

<table>
<thead>
<tr>
<th></th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>palato-alveolar</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stops p</td>
<td>t</td>
<td>t</td>
<td>k</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless ejective stops t'</td>
<td>t'</td>
<td>k'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced stop d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless affricate</td>
<td>ts, ts'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricatives</td>
<td>s</td>
<td>f</td>
<td>x</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasals</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximants w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>y</td>
</tr>
</tbody>
</table>

1.2.1 Phonology
Syllable structure is (C)V(C). Stress goes on the last syllable of a word. There are six vowels: i, e, a, o, u plus high central i. The consonants are set out in table 13.3. Notable features are the four ejective stops and affricates and the lateral fricative (these are not found in any other Xingu language). There is a language death phenomenon - some young speakers make no distinction between plain and ejective stops and affricates.

1.2.2 Features of the grammar
There are open classes of noun (including pronouns and demonstratives), verb, adjective and adverb and closed classes of auxiliary, postposition, number/quantifier, particle, subordinate clause marker and interjection.

Pronouns are analysable into root: ha (1sg and 1n.sg.excl), ka (1n.sg.incl), hi (2) or in-(3); plus zero (sg), a (dl) or wan (pl). 3sg has distinct masculine and feminine forms. Demonstratives 'this' and 'that' are based on 3rd person pronouns.

Trumai is a dependent-marking language. An NP in S and O function takes no marking and there are postpositions for ergative (A function), dative, locative, ablative, instrumental and comitative (the ergative, dative and locative markers are enclitics). Verbs fall into four classes: plain intransitive (obligatory S argument), plain transitive (A and O), extended intransitive (S and dative) and extended transitive (A, O and dative). Verbs in the extended intransitive class include 'eat', 'drink', 'see', 'like' and 'talk with'. There are two verbs 'kill' - fa which is extended intransitive and disi which is transitive.

Trumai has little morphology. There are no TAM suffixes to the verb, reference to time being expressed through adverbs. There are post-verbal particles marking.
the Portuguese in 1784 (partly, it is said, out of fear of the Mundurukú – see chapter 4 – an even fiercer people) but still indulged in raiding and killing into the nineteenth century. (Historical information from Rodrigues and Oliveira 1977, who list all sources.)

There were a number of dialects of Mura which appear (from the word lists collected – Martius 1867: 20-1; Nimuendajú and Valte-Bentes 1923: 218-21; Nimuendajú 1925: 160-6; 1932: 93-106) to have been fairly close. It seems that only one of these survives – Pirahã, spoken by about 100 people on the Maici river, a tributary of the Marmelos which itself flows into the Madeira. The Pirahã have retained their identity and language partly through maintaining an antipathy towards most aspects of the European-based culture that has surrounded them.

Three Summer Institute of Linguistics teams have worked on Pirahã – Arlo and Vi Heinrichs (from 1960 to 1966), Steven and Linda Sheldon (1967 until the late 1970s) and Daniel and Keren Everett (sporadically from 1978). They provide different accounts of several aspects of the language.

1.3.1 Phonology

Heinrichs (1964) put forward a system of eight consonants – see table 13.4 – together with three vowels (front i, central a and back o) and three tones (high 1, mid 2 and low 3). Everett states that vowels may optionally be nasalized following 7 or h.

Note that the allophones of /b/ include a bilabial nasal and also a bilabial trill. Heinrichs also described an unusual allophone of /g/ – a type of double flap in which the tongue tip hits the alveolar ridge and then (coming out of the mouth) the lower lip. (Later, D. Everett 1982 described these sounds, but without reference to Heinrich’s initial account.) D. Everett (1986) states that women articulate /h/ as [h], always before /l/ and sometimes elsewhere.

Sheldon (1974) accepted Heinrichs’ phonological analysis and described rules of tone assimilation (e.g., the first vowel of an adjective assimilates to the tone of the final vowel of a noun which immediately precedes it) and metathesis (aI→iIal-o, and oI→iIol-o, in certain circumstances). Everett (1986) states that he at one time recog-
Temporal clauses are marked by a suffix -sol-oo 'when, during, after'.

Negation is marked by verbal suffixes -sahal 'don’t’ and -hial ’not’. The latter also has a free form, hiala, used for negating NPs and nominalized clauses. There is a coordinator pitil ‘and, also’ (generally used with two clauses that have the same verb) and also hoagá ‘but’.

In a relative clause the argument in common with the main clause must be in a core (subject or object) function. Everett states that a relative clause can be marked
possessor; on a postposition, marking the postpositional argument; on an intransitive verb, marking the S, or on a transitive verb, marking the O. A number distinction is made only for 1st person: 1sg is 1, 1pl is hi, 2nd person is a, while 3rd person is i or e with nouns, postpositions and transitive verbs, but a with intransitive verbs (here falling together with 1sg). Free pronouns are 1sg hì, 1pl hìrì, 2 aje and 3 na.

Interestingly, a transitive verb must be preceded either by an NP (which can be a free pronoun) in O function, or by a pronominal prefix marking O. In contrast, every transitive verb bears a pronominal prefix indicating its A, whether or not there is an explicit NP in A function. A postposition is either preceded by an NP, or takes a pronominal prefix, e. g. a-paka ‘with you’, a-ri ‘for you’.

Post-verbal particles mark negation, tense, aspect and (at least with intransitive verbs) causative.

The constituent order is basically SV, AOV; A and S can be moved into postverbal position but O always immediately precedes the verb. Within an NP a demonstrative precedes and an adjective follows the head. An adjectival can function as NP head and must then take the 3rd person prefix i (unless nouns, verbs and postpositions, adjectives cannot take 1st and 2nd person prefixes).

As in many other Amazonian languages, nominalizations are employed where other languages would have complement or adverbial clauses. In the following examples, from Pires (1992: 132, 87), the post-verbal particle a marks the nominalization.

(22) hì A [a-hìi a hì] tôhe
1sg 2sgO-kill NMLZR want
‘I want to be your murderer.’ (i.e. ‘I want to murder you’)

(23) hì [a-rukàhì a ]
1sg 1sgS-arrive 2S-leave NMLZR
‘I arrived just as you left.’

It is likely that Jabuti originally had no numbers. In the contact situation ‘one’ is given as nici (we have no information about its original meaning). For ‘two’, je-ho is used, involving the verb root -ho ‘be equal’ and the reflexive-type prefix je-. For ‘more than two’ the phrase hìndì tò may be used, literally ‘know not’, i.e. ‘can’t say how many’. (Of course, Portuguese numbers are now used as loans, in addition.)

1.5 Chapacura family

Languages of the Chapacura (Txapacurun) language family, one of the smallest in southern Amazonia, are spoken in the Guaporé valley and along the tributaries of the Madeira river, in the west of the state of Rondônia and in the south of the state of Amazonas (Brazil), and in adjacent regions of Bolivia.

<table>
<thead>
<tr>
<th>Table 13.6 Wari' consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>voiceless stop</td>
</tr>
<tr>
<td>labialized stop</td>
</tr>
<tr>
<td>voiceless fricative</td>
</tr>
<tr>
<td>labIALIZED voiceless fricative</td>
</tr>
<tr>
<td>nasal</td>
</tr>
<tr>
<td>glottalized nasal</td>
</tr>
<tr>
<td>flap</td>
</tr>
<tr>
<td>glide</td>
</tr>
</tbody>
</table>

The genetic relationship of Chapacura languages has been known since d'Orbigny (1839); see also Créqui-Montfort and Rivet (1913) (the alternative name used for the family in earlier sources was Pawumwan). The first data on Torá go back to 1716 (Nimuendajú 1925, and see Nimuendajú and Valle-Bentes 1923); the first data on Urupá were collected at the end of the nineteenth century. The extinct languages of the family include Jarú (possibly a dialect of Urupá – Nimuendajú 1925: 139) and Chapacura (or Tapacura, also known as Huachis – Chamberlain 1910: 184). There are only short word lists in these languages. A superficial comparison of word lists shows that they are quite close (maybe as close as Tupi-Guarani, or as Romance). The only grammatical description is that of ‘Wari’ by Everett and Kern (1997), from which all information below is taken.

1.5.1 Phonology

The consonant system is given in table 13.6. There is also a complex sound made up of a voiceless apico-dental plosive followed by a voiceless labiodental trill (see Ladefoged and Everett 1996; Everett and Kern 1997: 1, 396). This occurs only before /l/ and /y/. It is reported to contrast with /l/ in one dialect (although only occurring in about twenty-five words) but is an allophone of /l/ in other dialects. There are six vowels, i, e, a, u, plus two front unrounded vowels, high open /i/ and mid-close /e/.

Syllable structure is CVC(C). There are elements of vowel harmony (not necessarily across morpheme boundaries; usually contiguous to r). Stress falls on the last syllable of the word (Everett and Kern 1997: 416).

1.5.2 Features of the grammar

All Chapacura languages are agglutinating with some degree of fusion, head-marking, with a nominative–accusative profile. The only open classes are nouns and
verbs; modifiers seem to be a subclass of nouns. There are no core case markers. A/S and O are cross-referenced through verbal inflectional markers (verbal clitics which contain reference to tense, aspect and mood). Cross-referencing is obligatory with definite objects, and optional with indefinite objects. Objects are not cross-referenced when frontal to sentence-initial position in relative clauses, information questions and clefted sentences.

In Wari', gender assignment is predominantly semantic with a certain degree of opacity (which is characteristic of languages of southern Amazonia). It goes along the following lines:

- feminine: human females, collective nouns, mixed groups of female and male individuals;
- masculine: human males, animals and culturally significant objects (natural phenomena, insects, fish, etc.);
- neuter: most inanimate objects, newly introduced objects/animals/plants; loans, nominalizations, mixed groups of neuter and masculine objects.

Typically for an Amazonian language, gender is not marked on the head noun. Gender agreement is marked on modifiers, possessed forms, and through verbal cross-referencing (Everett and Kern 1997: 294–300). A set of bound pronominal clitics cross-reference the person, number and possessor (if third person), e.g. caipi-in wao 'mouth-3neut basket' 'rim of the basket'; caipi-con won 'mouth-3masc cotton' 'hem of the skirt'; caipi-cam nairina (mouth-3fem woman) 'woman's mouth'; wijima-in xirim (smallness-3neut house) 'small house' (lit. 'the house's smallness'). The possessor is the head of an NP. There is one multi-purpose preposition which cross-references the person, number and gender of its argument.

Pronouns distinguish three persons, and three genders for 3rd person. There is singular/plural distinction in 1st and 2nd person and in 3sg.masc and 3sg.fem but not in 3sg.neut. 1pl distinguishes inclusive and exclusive. All these oppositions are maintained in reflexive/reciprocal pronouns.

Verbs distinguish reals and irrealis, marked with inflectional clitics which are portmanteau with person, number, gender, tense, aspect and mood. Tense distinctions are: recent past, past/present and future. Irrealis forms are also used as imperatives. Realis forms are used in prohibition. Clausal negation can be achieved either with the negative main verb 'om 'not exist', taking a complement clause, or by a postverbal modifier 'alara or ma'. Double negative meaning results in an emphatic positive reading. There are a few aspectual particles; serial verbs are used to express habitual aspect. Iterative aspect is marked by total reduplication of the verb root (Everett and Kern 1997: 323–8).

Some verbs have suppletive forms used when the S or O argument is plural, while other verbs mark plural (of S or O) by partial reduplication of the initial CV and infixation of -ra (1997: 337–9).

There is no morphological causative mechanism. There appears to be an (imperfective) passive; any verb can be passivized/intransitivized.

(24) wircam co quere see
EMPH:3sg.fem INFLECTIONAL:masc/fem REALIS.PAST/PRES
wa
PASSV
'She is the one who was seen.'

(25) 'irawin · con 'ac wa tara
afternoon sing travel PASSV 3sg.FUT.REALIS
'Then the afternoon will be sung in.'

Basic clausal constituent order is VXS, VOXS, where X is an indirect object. Verb compounding (or root serialization) is very productive. A typical directional compound is:

(26) pan' corom mama pin 'awi nana
fall enter go (pl.S) completely completely 3REALIS.PAST/PRES
'They all fell [fall-enter-go] into the water'.

Verb compounding is used to form causatives of switch-subject type on intransitive verbs only, e.g.:

(27) juc camara' na-in too Xijam
push fall (pl.S) 3sg.REALIS.PAST/PRES-3neut metal male.name
Xijam knocked down (push-fall) the cans.'

A causative of a transitive verb can be achieved by direct speech: 'he said: do it' for 'he made me do it'. Serial verb constructions are frequently lexicalized.

Subordination is marked on the verb. Relative clauses are marked with subordinating clitics. Only core arguments can be relativized. Nominalization is also used to mark subordinate clauses.

1.6 Mákú2

This language isolate had only three speakers in 1964, living near the Uraricuera river, Roraima, northern Brazil. The Mákú, according to their own stories, originally lived around the Maluwaka mountains between the upper Padamo and

2 This should not be confused with the Mákú family -- with different stress -- discussed in chapter 9.
Aikaná. Older sources - Hanke (1956) are preliminary studies Aikaná (also known as Kwaza, Koaia). Their migration from Venezuela to Brazil was caused by constant attacks by Yanomama. Numbers fell drastically in the period from the 1930s to the 1950s, due to diseases. The Máku were first mentioned by Koch-Grünberg (1913: 458). A short description of Máku phonology and elements of morphosyntax are found in Migliazza (1965, 1966 and 1978a – which also contains a short word list), unfortunately cast in opaque tagmemic formulas.

The consonants of Máku are in table 13.7 (Migliazza 1978a). There are four vowels (i, e, a, u) existing in both oral and nasal varieties, plus two additional oral vowels, high front rounded i and high-central e. Stress is not contrastive (its typical position is on the last syllable); vowel length is contrastive but occurs only in an initial C(V) syllable in a polysyllabic word. Syllable structure is (C)(CV(C).

Máku is head-marking, highly polysynthetic and predominantly suffixing. There are no genders or classifiers. Pronouns distinguish first person exclusive and inclusive. There is a very complex system of tense-aspect oppositions.

1.7 Aikaná and Koaia

Aikaná (also known as Tubarão, Huari, Masaká, Kasupá and Mundé) is spoken in southeastern Rondônia by about 120 people including some children (Hein van der Voort, p.c.).

There is only a little information available – fieldnotes by Harvey Carlson (1986) (Hinton 1993 is based on this), and papers by Vasconcelos (1993a,b, 1996) which are preliminary studies of the phonology and some aspects of morphology of Aikaná. Older sources – Hanke (1956) and Becker-Donner (1955) – contain just a few words.

The consonant system, in table 13.8, includes a dental fricative (a rarity in the Amazon). There are oral and nasal forms of four vowels, i, e, a, u plus a fifth oral vowel, front rounded ù. It is unclear whether there are contrastive tones.

Aikaná is predominantly suffixing with some prefixes. There is no inclusive/exclusive distinction in pronouns. There are two genders or noun classes; animacy is distinguished in interrogatives (bari 'who', tara 'what'). Most adjectival meanings are expressed through verbs; colour concepts are expressed by nouns. Possession is marked with suffixes, and with genitive forms of personal pronouns; these can co-occur. Grammatical relations are marked by cross-referencing on the verb: subject (A/S) is marked with suffixes or a combination of prefixes and suffixes, while object is marked with suffixes. There is an accusative case on NPs. There is body-part incorporation, verbal classifiers, and possibly incorporating serial verbs. Verbal classifiers appear in the first position in the verbal word; they refer to the S, or the O, characterizing it in terms of its nature ('liquid', 'fibrous', 'metallic'), or its structure ('powder-like', 'consisting of coarse particles', e.g. seeds or kernels); there is one classifier covering animates and humans. Initial CV duplication indicates repetitive action. There is probably a benefactive valency-increasing derivation. The verb is clause-final and there are postpositions.

Koaia (also known as Kwaza) is spoken by about 25 people who live together with the Aikaná. The only information available comes from preliminary studies by van der Voort (1997a, b) (a short word list is given in Loukotka 1963).

The consonant system is given in table 13.9. There are oral and nasal variants of seven vowels, i, e, a, i, a, o, u. Syllable structure is CV(V). It is unclear whether stress (on the last syllable of the stem) is contrastive. Glides w and y are allophones of u and i.

Koaia appears to be predominantly suffixing and combines dependent-marking with head-marking. The verb takes an obligatory pronominal suffix for S and a further suffix, which appears to be optional, for O. Only 1st and 2nd person dis-
2.1 The Tacana family

The Tacana family consists of four or five living languages spoken mostly in northwest Bolivia and in the adjacent areas of Peru, between the rivers Madre de Dios and Beni (table 13.1). Loukotka (1968: 175–6) and Girard (1971: 20; based on Mason 1950) mention a few more languages of the Tacana family, all of them extinct.

Table 13.9 Koai consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>alveolar</th>
<th>apico-alveo-palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>?</td>
</tr>
<tr>
<td>implosive voiced stop</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td>k*</td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>s</td>
<td>f</td>
<td>h</td>
<td>m</td>
<td>n</td>
</tr>
<tr>
<td>approximant</td>
<td>ts</td>
<td>tf</td>
<td>l</td>
<td></td>
<td>r</td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

tinquish singular and plural; there is also an inclusive/exclusive distinction. There appear to be verbal classifiers; the same morphemes are used as agreement markers on nominal modifiers. There is a complex system of tenses (immediate past, proximate past, remote past) and aspects, and also directional markers on the verb. Conditional and concessive are also marked on the verb. Clausal constituent order appears to be free.

2 BOLIVIA

There are about 35 indigenous languages spoken in Bolivia by around 70,000 people. Besides languages of the major families – Arawak, Pano, Tupí-Guaraní, as well as Quechua and Aymara – languages spoken exclusively, or almost exclusively, in Bolivia include those belonging to the Tacana family (§2.1) and a number of isolates (§2.2). Here we will consider only languages spoken in the northern parts of Bolivia, which fall within Amazonia.

2.1.1 Phonology

Cavineña (Key 1968: 19) has the largest phonological system – see Table 13.10. It includes a series of alveo-palatal phonemes missing from other Tacana languages. Tacana also has interdental voiced stop d, while Ese Eja lacks ts. According to Key (1979: 85), Ese Eja has a series of implosive stops; there is also considerable allophonic variation, e.g. /n/ can be realized as [nd], [l], [nl] or [d]. All Tacana languages have four vowels, i, e, a, o (with a as an allophone).

Stress is not contrastive (it falls on the penultimate syllable of a root in Cavineña and Tacana). Ese Eja has a complicated system of stress assignment which is

Table 13.10 Cavineña consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td>k*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>labialized stop</td>
<td>m</td>
<td>n</td>
<td>j</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>s</td>
<td>f</td>
<td>h</td>
<td>m</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>voiceless approximant</td>
<td>ts</td>
<td>tf</td>
<td>l</td>
<td>m</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>j</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
different for nouns and verbs. The common syllable pattern is CV (VC structures are found in Tacana and only rarely in non-loans in Cavineña).

Araona has a male–female speech distinction (/in female speech corresponds to s in male).

2.1.2 Features of the grammar
All Tacana languages are head-marking and polysynthetic, with long strings of verbal suffixes to express tense, manner of action and modality. They mostly use suffixes; a few prefixes occur only with verbs (Key 1968: 29). There are no genders, or classifiers. Valency-changing derivations are marked by verbal prefixes and suffixes. Tacana languages appear to have auxiliary verbs. Cases are used for marking grammatical relations on all types of NPs in Tacana and Cavineña; in Araona and Ese Eja case markers are apparently used only with pronouns. Basic clausal constituent order is SV, AVO.

At least some Tacana languages show ergative properties. Cavineña has a very unusual pattern of split ergative marking of grammatical relations with cases. It combines split-ergativity conditioned by (a) the semantics of nouns, (b) mood and polarity, and (c) whether a clause is main or subordinate (Camp 1985).

Ergative (-ra) is obligatory with nouns in A function, but may be omitted from a pronoun in A function. 'If the overt noun is ergative, then the absolute pronoun is the object [...] if the noun is absolutive, then the absolute pronoun is the subject' (Camp 1985: 44). In (28) both the pronominal A ('we two') and the direct object ('cows') are marked for absolute case. In (29) the non-pronominal A ('cows') is marked with ergative case.

(28) Tuja ya-tse waka k’ana ba-tse-tjine umada so 1-DEL+ABS cow PL+ABS see-arriving.O-PAST many 'So we two saw many cows coming'

(29) A-ya ya-tse iyak’a waka-ra do-PRES 1-DEL+ABS now cow-ERG 'The cows will get us now.'

Ergative marking is obligatory with pronouns if the sentence is negative, potential, counterfactual or intentional. Example (30) illustrates the ergative marking on a pronominal A in a negative clause.

(30) pake me-tse-ra e-k’e utsek’a nuru-tjine ama apparently 2-DL-ERG 1-GEN grandchild+ABS care.for-PAST not 'Apparently you two didn’t take good care of my grandchild.'

---

13 Other small families and isolates

<table>
<thead>
<tr>
<th>Table 13.11 Cayuvava consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>voiceless stop</td>
</tr>
<tr>
<td>voiced stop</td>
</tr>
<tr>
<td>voiceless fricitive</td>
</tr>
<tr>
<td>voiced fricitive</td>
</tr>
<tr>
<td>nasal</td>
</tr>
<tr>
<td>flap</td>
</tr>
<tr>
<td>glide</td>
</tr>
</tbody>
</table>

Ergative marking on personal pronouns is also obligatory in subordinate clauses, as shown in (31).

(31) i-ke ni mi-ra ara-kara a-ya, duhu-k’e 1-affix + ABS if 2-ERG eat-want do-present take-IMPER
  i-ke espere hiruru 1-affix + ABS stream edge
  'If you want to eat me, take me to the edge of the stream.'

Personal pronouns are marked for ergative case if they appear sentence-initially, as in (32), or sentence-finally (in fact, this is encountered only rarely):

(32) e-ra bak’e a-k’are 1-ERG contrast do-REM.PAST
  'I, A said'

Cavineña is unusual in that it appears to have an antipassive-like valency-reducing derivation, marked with reduplication of the verb stem; then A becomes S (i.e. is marked with absolutive case), and O is omitted.

2.2 Isolates

A list of isolates spoken in northern and northeastern Bolivia is given in table 13.1. The materials available for these languages are of mixed quality. For Cayuvava there is a short grammar by Harold Key (1967); for Itonama there is a short grammar and a dictionary by Camp and Liccardi (1965, 1967), and for Movima there is a short grammar and a dictionary by Judy and Judy (1962, 1965). Unfortunately, both grammars are executed in a non-user-friendly tagmemic model.

The Bolivian isolates have rather diverse phonological systems. Syllable structure is (C)V. The consonants of Cayuvava are in table 13.11 (H. Key 1961; 1967;
Table 13.12 *Itonama* consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>ty</td>
<td>k</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glottalized stop</td>
<td>t'</td>
<td></td>
<td></td>
<td>k'</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>β</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glottalized affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>

Table 13.13 *Movima* consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>dental</th>
<th>alveo-palatal</th>
<th>palatal</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>labialized stop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>preglottalized nasal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>

15). Cayuvava has $i$, $e$, $a$, $o$, $u$ plus high central $i$, open-mid front $o$ and mid-close back $a$.

*Itonama* (Liccardi and Grimes 1968; Camp and Liccardi 1965) is unusual in that it has a series of glottalized stops, and also palatalized $t$ — see table 13.12. There is a typical Amazonian system of six vowels: $i$, $e$, $a$, $o$, $u$ plus high central $i$.

The consonantal system of *Movima* is less complex (Judy and Judy 1965) — see table 13.13. It is unusual in having preglottalized nasals. There are just five vowels $i$, $e$, $a$, $o$ and $u$.

Cayuvava has contrastive nasalization of vowels and contrastive stress. Nasalization, in addition, is reported to function as a kind of honorific device: a person of lower social or economic status addresses another one of a higher rank with a prominence of nasalization on all the vowels of the utterance (Key 1979: 83).

All the Bolivian isolates are head-marking, polysynthetic and nominative-accusative. Cayuvava has a most complicated morphology. Its striking property is the existence of six prefix positions on verbs — these include passive–reflexive, thematic, cross-referencing of person–number of the subject, tense and aspect. Nouns take derivational prefixes, and discontinuous circumfixes for marking possession. There are no genders or classifiers. There are only two open classes — nouns and verbs; a nominal modifier can be derived from any word class. Constituent order is apparently free.

*Movima* appears to be predominantly suffixing with just one prefix position (occupied by the cross-referencing pronoun). In *Movima*, suffixes are used for causative, benefactive, malefactive, reflexive and a number of tense–aspect, modality and directional categories (e.g., Judy and Judy 1965: 188–9). Itonama has prefixes, suffixes and infixes (1965: 293); besides cross-referencing, many other categories — such as causative, benefactive, directional — are marked with prefixes to the verb.

An unusual property of *Movima* is its system of numeral classifiers. It also has three genders (masculine, feminine, neuter or inanimate) in personal pronouns and cross-referencing (1965: 202). Preliminary fieldwork by Grinevald (1996) shows that the classifiers include *-poy* 'quadruped animals', e.g. fox, crocodile, tapir, etc.; *-mo* 'biped animals', e.g. rooster, duck, owl, etc.; and *-ba* 'fruit', e.g. papaya, guava, orange (cf. also Key 1979: 67–8). For other native *Itonama* nouns, the last syllable is repeated on the numeral as an agreement device, e.g. *-d'o* for *chad'o* 'plate'; *-mas for *d'imas* 'hay'; *-pi, for sukapi* 'belt'. For borrowed nouns, the last two syllables are repeated if a noun consists of more than two syllables, e.g. *-mis'a* for *kamisa* 'shirt', and *-pato* for *zapato* 'shoe' (both loans are from Spanish). If a loan consists of just two syllables, the reduplicated last syllable is repeated, e.g. *-sasa for mesa* 'table', *-yaya for siya* 'seat, chair'.

*Itonama* distinguishes masculine and feminine genders in first, second and third person singular pronouns and in demonstratives (Camp and Liccardi 1965: 331–2; Key 1979: 40). It has a very complex system of at least seventeen classifier suffixes based on animacy, shape and position of object; classifiers combine with adjectives, verbs and demonstratives. *Itonama* may also have body-part incorporation (Key 1979: 41).

3 COLOMBIA

There are about 66 indigenous languages in Colombia. They belong to 22 language families and are spoken by approximately 500,000 people (Landaburu 1994b). There are also two creoles of Afro-European origin in north-west Colombia.
addition, Lingua Geral Amazônica (Nheengatá), a creole of Tupí-Guaraní origin, is spoken on the border with Brazil.

Languages from the larger families are discussed elsewhere in this volume – Carib in chapter 2, Arawak in chapter 3, Tupí-Guaraní in chapter 5, Tucano in chapter 7, Makú in chapter 9, and Witoto plus the isolate Ticuna in chapter 12. Languages of the Guahibo and Sáliba–Piaroa families and the isolate, Andoké, are spoken exclusively in Colombia. Their typological properties are briefly discussed here.6

The majority of the indigenous peoples of Colombia are slash-and-burn agriculturalists. Only the Guahibo peoples were nomadic hunters and gatherers, like the Makú peoples in Colombia and Brazil (see chapter 9). Guahibo constitutes an unusual example of a large nomadic community.

Guahibo and Sáliba are currently being studied by scholars from the Centro Colombiano de Estudios en Lenguas Aborígenes in the Universidad de los Andes in Bogotá, and by members of the SIL (see Queixalos forthcoming, and n.d., on Sikuani/Guahibo; Kerr 1995 on Cuiba; Estrada 1996, and Morse and Frank 1997, on Sáliba). There is a large but opaque grammar of Andoké by Landaburu (1979), which also contains a few texts. There is virtually no grammatical information on Piaroa. Huber and Reed (1992) give an overview of phonological systems of most Colombian languages, accompanied by 200 words in each language.

3.1 Phonology

The Guahibo languages, Sáliba and Andoké have a voice distinction in stops. There is typically one liquid with a rhotic allophone. Cuiba is typologically unusual in that it has no rhotic or lateral; there is, however, a trilled r as an allophone of d (Kerr 1995: 19). This phonological oddity may have an areal explanation.7 Sáliba has two rhotics and a lateral.

4 More evidence is needed to prove the genetic link between Sáliba and Piaroa: see Landaburu (1994b: 370) and Estrada (1996: xxv–xxvi).

5 There is no adequate information on Puinave which is possibly another isolate (although it has been suggested that it may be distantly related to Makú).

6 We do not discuss, in this volume, language families whose members mostly or entirely fall outside Amazonia – Chibchan, Paez, Choco, Barbacoan, Kuna (all of which have – almost certainly mistakenly – been grouped together as ‘Macro-Chibchan’).

7 The dental stop is realized as a lateral flap in north-west Amazonia (Tucano-Tariana region) and also further to the south-west. r is an allophone of d in proto-Witoto (Aschmann 1993: 96). Resigare, a language of the Arawak family, which came under strong areal influence from Witoto languages, has r; in accordance with Allin (1975), it may have once been an allophone of d. Other r-less languages, besides Cuiba, are concentrated to the west of the Vaupés; Huprua (Makú), and Aguaruna in the adjacent regions of Peru (there is here a lateral fricative, f, which is an allophone of r before h). Allophonic variation between a dental stop and a rhotic stretches further to the west (it is also attested in the Barbacoan language Awa Pit: Curnow 1997).

---

### Table 13.14 Cuiba consonants

<table>
<thead>
<tr>
<th>Type</th>
<th>Bilabial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>h</td>
<td>?</td>
</tr>
<tr>
<td>aspirated stop</td>
<td>pʰ</td>
<td>tʰ</td>
<td>kʰ</td>
<td>hʰ</td>
<td>?</td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>s</td>
<td>tʃ</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>voiceless affricate</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Table 13.15 Sáliba consonants

<table>
<thead>
<tr>
<th>Type</th>
<th>Bilabial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>-</td>
<td>?</td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>lateral</td>
<td>l</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>glide</td>
<td>w</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The consonant systems of Cuiba (Guahibo: Huber and Reed 1992; Kerr 1995), Sáliba and Andoké are given in tables 13.14–13.16.

Most Guahibo languages have six vowels (i, e, a, o, a plus high central ĭ); Sáliba has just five, lacking ĭ; in each language there is a nasal vowel corresponding to each oral vowel. Andoké has a larger system with nine oral and five nasal vowels, as in table 13.17 (Landaburu 1979: 45–6).

Typical syllable structure is (C)V; Cuiba and Sáliba can have a nasal in coda position, and Cuiba also permits stops.

The Guahibo languages and Sáliba have contrastive stress (Estrada 1996: 3). In Guahibo, stressed vowels tend to have high pitch (Kondo and Kondo 1967: 95). Andoké appears to have three tonal distinctions: high, middle and low (Landaburu 1979: 48–51). This property is shared with two neighbouring Witoto languages, Bora and Ocaina, which distinguish two levels of tone (see chapter 12).
Table 13.16 Andoké consonants

<table>
<thead>
<tr>
<th></th>
<th>bilabial</th>
<th>alveolar</th>
<th>velar</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless stop</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>?</td>
</tr>
<tr>
<td>voiced stop</td>
<td>b</td>
<td>d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless fricative</td>
<td>f</td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced affricate</td>
<td>dʒ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glide</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 13.17 Andoké vowels

<table>
<thead>
<tr>
<th>oral</th>
<th>nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>front</td>
</tr>
<tr>
<td></td>
<td>front</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>e</td>
<td>ø</td>
</tr>
<tr>
<td>a</td>
<td>ø</td>
</tr>
<tr>
<td>ü</td>
<td>ø</td>
</tr>
</tbody>
</table>

3.2 Morphology

3.2.1 Word structure and typological profile

All of these languages combine head-marking and dependent-marking properties. They are agglutinating with polysynthetic tendencies, the Guahibo languages being the most polysynthetic of all. Typically for North Amazonian languages, there is one prefix position and numerous suffix positions. Some Guahibo languages are unusual in that nouns have more extensive suffix possibilities (up to six in Cuiba – Kerr 1995: 267) than verbs. The prefix position on a verb is occupied by a cross-referencing pronoun, valency-increasing marker, or by a marker of plural for S or O argument. Interestingly, in Sáliba verbal cross-referencing affixes are prefixes only on vowel-initial roots, while for consonant-initial roots they are infixes or the first suffix (Paul Frank p.c.). In Andoké the cross-referencing pronouns are all suffixes (Landaburu 1979: 130). In Guahibo, nouns may take a possessive prefix.

The Guahibo languages are unusual for northern Amazonia in that they have suppletive forms of verbs relating to the number of the S or O argument, e.g. Cuiba bêxuba ‘kill (sg O)’; jutebá ‘kill (many O)’.

Adjectives are an open class, except in Guahibo where they are said to be a small class of about twenty members (Francisco Queixalos p.c.).

3.2.2 Nominal morphology

Free pronouns and cross-referencing pronominals distinguish three persons and, generally, two numbers. Guahibo languages are unusual in that they distinguish singular, dual, pronominal and distributive plural, and also first person inclusive and exclusive (Francisco Queixalos p.c.). In some Guahibo languages (e.g. Cuiba), there are dual forms of nouns, tonic pronouns and verbs, but not of cross-referencing bound pronouns; in other languages (Guayabero: Tobar Ortiz 1994: 516–17), dual is marked on all types of nouns, pronouns and verbs.

All the languages discussed here have gender in third person pronouns, plus a large and complicated system of classifiers. Guahibo languages distinguish three genders (feminine, masculine, inanimate) in independent 3rd person singular and dual pronouns, but these are neutralized in the plural. There are no gender distinctions in bound pronouns. Sábiba is reported to mark three genders on 3rd person pronouns and adjectives: feminine, masculine and inanimate (Suarez 1977: 21).

Guahibo languages have complex systems of classifiers. The same morphemes are used as noun class agreement markers on adjectives, as in (33); with numerals, as in (34); and with deictics, as in (35). Adjectives, numerals and deictics with classifiers can be used headlessly. Examples (33)–(35) are from Cuiba (Kerr 1995: 243ff.).

(33)    peru-nae
        old-CL:WOODEN THINGS
        ‘old (canoe)’

(34)    cae-bo
        one-CL:HOUSE
        ‘one (house)’

(35)    barapo-bo
        this-CL:HOUSE
        ‘this (house)’

In Cuiba possessive constructions, classifiers are used independently of whether the noun is alienably or inalienably possessed, just like in neighbouring Tucano and North Arawak languages of the Upper Rio Negro region:

(36)    (Kerr 1995: 134)
        piya-nae jera
        3sg.POSV-CL:WOODEN THINGS canoe
        ‘his/her canoe’

Also as in East Tucano languages and Tariana, classifiers in Guahibo can be considered an open class due to the existence of repeaters (nouns used to classify them-
selves); however, only inalienably possessed nouns can be used as repeaters (1995: 132).

Sáliba also has a system of at least thirty classifiers. The same set of morphemes goes on numbers, adjectives, demonstratives, possessives and even interrogative pronouns. Some classes distinguish a singular and a plural form. Interestingly, numbers 'one' and 'two' also have different forms for animate and inanimate. (Estrella 1996: 106-7).

In Guahibo languages and in Sáliba (Paul Frank p.c.), the classifier suffixes used with nouns have an individuating effect: a noun is unmarked for number if used without a classifier. If a classifier is added, the noun has a singulative reading and can additionally be marked for number, e.g. Cuiba (Kerr 1995: 154) asocom ‘wild fruit’, tsocoma-re (fruit-CL:FRUIT) ‘(one) wild fruit’, tsocoma-re-a (fruit-CL:FRUIT-PL) '(several) wild fruits’. This individuating effect of noun classifiers is reminiscent of Tucano languages and Tariana of the Vaupés area, and of Resigaro and Witoto-Bora. It may be the areal feature of a region embracing northeastern Peru, south-central Colombia and north-western Brazil.

Andoké also has a complex agreement system. Verbal cross-referencing markers have three noun classes. Class 3 divides into four subclasses: cross-referenced on the verb in the same way, but marked differently on pronouns and noun modifiers. The class assignment is far from semantically transparent. The classes are:

(i) Class 1 includes inanimate objects, mostly artifacts, but also some animals, some mythical beings, and plants.
(ii) Class 2 includes mainly long and hard objects.
(iii) Class 3 includes most animates and also round and hollow objects. Within this class, the following subdivisions are made:
— Class 3a contains round and hollow objects;
— Class 3b contains masculine animates;
— Class 3c contains feminine animates;
— Class 3d contains animate collectives.

Guahibo, Sáliba and Andoké have an unusually large number of oblique cases compared to other Amazonian languages (such as Tucano). There is one ‘multipurpose’ oblique case (e.g. Cuiba -tha ‘locative, directional, comitative’, or Andoké -a ‘locative, indirect object marker’), and a number of other cases with more specific meanings, e.g. Cuiba -xar ‘because of’, -xaña ‘because of (absence of)’, -nexa ‘purpose’, -yainwa ‘for fear of’ (Kerr 1995: 107, 170); Andoké -mä ‘comitative, instrumental’, -ta ‘private’, -tao ‘benefactive’, -a ‘malefactive (to the detriment of)’ (Landaburu 1979: 166-7); Sáliba -di ‘dative’, -da ‘ablative-allative’, -na ‘inessive’, -gi ‘comitative’, -digi ‘comitative’ (Estrella 1996: 92-5).

3.2.3 Grammatical relations
Grammatical relations are marked by cross-referencing. Only Guahibo languages have some traces of split ergativity, of an active-stative type.

In Cuiba, verbs divide into five conjugations which take different tense markers and show different vocalic alternations (Kerr 1995: 31ff.). Different cross-referencing suffixes are used for S1, S2, and S3 of positive and negative clauses (Guayabero – Tobar Ortiz 1994: 521–2; Cuiba – Kerr 1995: 31). Another set of cross-referencing suffixes is used for the majority of stative S verbs (Kerr 1995: 59–60). A special set of prefixes are used for O, and also for S0 of some verbs which refer to physical states, e.g. Cuiba:

(37) (Kerr 1995: 51)
\[ \text{xam ne-jë-rate-me} \]
\[ \text{you 1sg + O-look.for-2sg + A} \]
‘You looked for me.’

(38) (Kerr 1995: 71)
\[ \text{ne-suyaba} \]
\[ \text{1sg + S_have.diarrhoea} \]
‘I have diarrhoea.’

3.2.4 Verbal morphology
There are usually several valency-increasing verbal derivations. In Guahibo the prefix ku-functions as an applicative and is also used to form causatives of transitive verbs, while any verb can be causativized with an auxiliary (Francisco Queixalos p.c.). In Guahibo languages the same morphemes can be used as applicative markers on verbs, and as case-like markers on pronouns and inalienably possessed nouns. A privative applicative verbal prefix is illustrated in (39), from Cuiba (Kerr 1995: 107):

(39) ta-xant-yo tsi-se-ta-n
\[ \text{1sg-child-FEM PRIVATIVE.APPL;COOK-1sg-PAST} \]
‘I cooked for my daughter without her help.’

If the NP promoted to be a core argument is a pronoun or an inalienably possessed noun, the applicative marker is suffixed to the NP, as in (40) (1995: 107):

(40) ne-tsi xane
\[ \text{1sg-PRIVATE,APPL_{-}3sg + eat} \]
‘He ate without me.’

Andoké has two causatives, and also a benefactive applicative (Landaburu 1979: 204-5).
There are also a number of valency-decreasing derivations; Guahibo has passive, middle, and a curious derivation whereby the addressee of a ditransitive verb becomes an oblique, if human (Francisco Queixalos p.c.). In Cuiba the reflexive-reciprocal derivation can also be used to emphasize the identity of the agent, similar to English self in I did it myself. Andoké appears to have just one general intransitivizer (Landaburu 1979: 205).

All the languages have rich systems of mood and modality, as well as directional suffixes on verbs. Guahibo languages have a restricted system of evidentiality (reported vs non-reported).

Guahibo languages have unusual incorporation patterns. Typically for Amazonian languages (e.g. Weir 1990, on Nadèb), only inalienably possessed nouns can be incorporated. Cuiba has two patterns of noun incorporation:

(i) inalienably possessed nouns, in S function, can be infixed into verbs of physical state (Kerr 1995: 104), e.g. atane 'hurt', a-cobe-tane 'his hand hurts'.
(ii) with other verbs, an NP in S, O or an oblique function can be incorporated (but not infixed) into a verb. Typically for Amazonian languages, incorporated nouns are preposed to the verbal root. Unlike most other Amazonian languages — but like Nadèb (Makú family) — Cuiba permits incorporation of more than one (up to three) arguments (Kerr 1995: 278–9); here incorporation does not affect the transitivity of a verb, e.g. (41) and (42). However, in Guahibo noun incorporation is a valency-reducing derivation (Francisco Queixalos p.c.).

(41) (Kerr 1995: 280)
  cobe-fifina-n
  hand-tired-1sg
  'My hand is tired' or 'I am hand-tired.'

(42) (Kerr 1995: 279)
  na-mali-péré-na-dobóba-me
  refl-arm-skin-hair-take off-2sg
  'You take off the hair of the skin of the arm.'

3.3 Syntax

Causal constituent order tends to be free. There is no verb serialization. Guahibo languages have some verb compounding (Kerr 1995: 171). A reduplicated main verb plus pona 'go' has an intensive meaning (1995: 215).

Guahibo languages mark relative and complement clauses by nominalizations (1995: 119). Cuiba also has full relative clauses marked by a relative pronoun (1995: 247). Guayabero has a relative clause marker which goes on the common argument within the relative clause if it is also in subject function in the main clause. Guahibo and Andoké have a large number of suffixes which mark the predicate of a subordinate clause. Subordinating suffixes often indicate the temporal relationship between main and subordinate clause (simultaneous, preceding or following). These languages do not have switch-reference constructions.

4 VENEZUELA

In Venezuela there are about 38 indigenous languages spoken by around 60,000 people. Languages from the larger families are treated elsewhere in this volume – Carib in chapter 2, and Arawak in chapter 3. Língua Geral Amazônica, of Tupí-Guarani origin, is spoken in the south. There are also three isolates – see table 13.1. The Yaruro, Warao and Hoti are basically monolingual nomadic hunters and gatherers. However, according to Mitrani (1988: 164) and Heinen (1988: 611), the Yaruro and Warao have for some time practised slash-and-burn agriculture and in the 1930s the Warao commenced rice cultivation.

There are a few papers on Yaruro (Obregón Muñoz 1981, and references therein). For Warao, there are a number of papers by Osborn (1966a,b, 1967), by Romero-Figueiroa (1986a,b), and a short grammar by Romero-Figueiroa (1997). There are almost no linguistic materials on Hoti (only a word list in Mattéi-Müller, Reid and Henley, 1994).

4.1 Phonology

Yaruro has a large phonemic system – see tables 13.18 and 13.19 (Obregón Muñoz 1981). It is unusual for the region in having a velar fricative (as well as a glottal fricative) and a velar nasal, in having voice distinction for all the stops, and in having some fricatives.

The consonant system in Warao is much simpler, as set out in table 13.20 (Romero-Figueiroa 1997: 105). There are no voice distinctions in stops or fricatives ([d] is an allophone of [r] in word-initial position). The only unusual phoneme is a labialized velar stop. There are five vowels (i, e, a, u, o), all with nasal counterparts (Osborn 1966a: 109). There is no contrastive length. Syllable structure is (C)V. Stress is on the penultimate syllable of a word.
There is a special verbal suffix indicating singular or plural number of the subject (S or A) argument. Possessive prefixes on nouns and object cross-referencing prefixes on verbs are almost identical. There are four oblique cases (dative, locative, allative and ablative). There are also a number of postpositions (causal, comitative, instrumental). There are no genders or classifiers. Warao has an agentive passive marked with completive past tense and perfective aspect on the verb; the demoted A may be introduced with the agentive postposition; compare the active clause in (43) with the corresponding passive in (44) (Romero-Figueroa 1997: 93).

(43) rihawaratuma\textsubscript{\textit{A}} \{ka-ina kokotuka\textsubscript{\textit{b}}\}, non-a-e ancestors 1pl.poss-land all make-PUNCTUAL-PAST 'Our ancestors made all in our land.'

(44) \{ka-ina kokotuka\textsubscript{\textit{b}}\} \{rihawaratuma aisiia\textsubscript{\textit{b}}\}, 1pl.poss-land all ancestors AG nona-i-ha make-COMPL-PERFV 'All in our land was made by [our] ancestors.'

Causative is marked with a prefix (\textit{e-}) on transitive and intransitive verbs. When a transitive verb is causativized, the underlying A of the causativized verb becomes the O of the new causative, while the underlying O becomes oblique (e.g. acquires dative marking) (Romero-Figueroa 1997: 94).

(45) wahabu-ma bare-tirma\textsubscript{\textit{A}} \{ka e-nahoro-a-e\} venison-DATIVE father-FEMALE 1pl.O CAUS-eat-PUNCTUAL-PAST 'The nuns made us eat venison.'

There is no verb serialization; an auxiliary verb is used to express certain mood meanings (e.g. prohibitive, potential and negative potential, and interrogative). Subordination is marked with suffixes on the verb and with particles.

Warao has an elaborate system of speech styles: 'council style', 'legend style' and usual style. Besides lexical differences, the 'council style' differs from other styles in more frequent use of some morphological characteristics, e.g. suffixal pronouns and morphological causatives of ditransitive verbs.

4.2 Features of the grammar

Warao is head-marking, mildly polysynthetic and nominative-accusative. Its basic constituent order is SV, OAV. There is no clear-cut distinction between nouns and adjectives. It is predominantly suffixing, with few prefixes. Direct object cross-referencing markers are prefixed to the verb, while subject cross-referencing markers, plus tense-aspect, modality and directional morphemes, are suffixes.

<table>
<thead>
<tr>
<th>Table 13.18 Yaruro consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>voiceless stop p</td>
</tr>
<tr>
<td>voiced stop b</td>
</tr>
<tr>
<td>voiceless affricate</td>
</tr>
<tr>
<td>voiceless fricative f</td>
</tr>
<tr>
<td>voiced fricative v</td>
</tr>
<tr>
<td>nasal</td>
</tr>
<tr>
<td>flap</td>
</tr>
<tr>
<td>lateral</td>
</tr>
<tr>
<td>glide</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 13.19 Yaruro vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>ae</td>
</tr>
<tr>
<td>a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 13.20 Warao consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>voiceless stop p</td>
</tr>
<tr>
<td>labialized stop s</td>
</tr>
<tr>
<td>voiceless fricative f</td>
</tr>
<tr>
<td>nasal</td>
</tr>
<tr>
<td>flap</td>
</tr>
<tr>
<td>glide</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY**


1968. Classification of South American Indian languages. Los Angeles: Latin American Center, UCLA.


1966. Esebo sintáctico de um corpus da língua Maku, Boletín del Museo Paraguayo Emilio Goeldi, Antropología, 32.


Obregón Muñoz, H. 1981. 'En torno al alfabeto de las lenguas indígenas de Venezuela a propósito de la reducción del pumí a la escritura', Revista Latinoamericana de Estudios Etnolingüísticos, 1, 25-46.


1966b. 'Warao II: nouns, relatives and demonstratives', IJAL, 32, 253-61.

1967. 'Warao: verbs and suffixes', IJAL, 33, 46-64.


13 Other small families and isolates


Schuller, R. 1933. 'The language of the Tacana Indians (Bolivia)', Anthrobus, 28, 99-116, 463-84.


Vasconcelos, I. 1933a. 'Marcador de transitividad en Aikana', paper at Workshop on Grammatical Categories, Campinas, Universidade Estadual de Campinas.

1933b. 'El aikana phonology: a preliminary analysis', Ms.


Voort, H. van der. 1997. 'Enkele problemen in de analyse van het Kwaza (Koaid)', a seminar at the Institute for Functional Research into Language and Language Use, Amsterdam.


Areal diffusion and language contact in the Ñana-Vaupés basin, north-west Amazonia

ALEXANDRA Y. AIKHENVALD

1 AREAL DIFFUSION IN NORTH-WEST AMAZONIA

To define a linguistic area, a convincing number of diagnostic traits must be established which will distinguish this area from others, and which can be proved not to be genetically inherited. Distinguishing areal diffusion from genetically inherited phenomena is particularly important for Amazonian languages—see chapter 1 above.

The region around the Upper Rio Negro (the major northern tributary of the Amazon) contains a multiplicity of languages. Within the Upper Rio Negro area, the combined Vaupés and Ñana river basins constitute a well-defined linguistic area; and within this the Vaupés basin is a clear linguistic sub-area (AIKHENVALD 1996a).

I shall here consider the Vaupés-Ñana basin in north-west Amazonia as an example of a linguistic area, which displays a convincing number of structural features shared by languages of three genetically unrelated families—North Arawak, East Tucano and Makú. These features are not found in Arawak and Makú languages spoken outside the area, and thus can be considered as diagnostic for areal diffusion. In some cases we are able to establish the direction of diffusion.

The languages spoken in the Ñana and Vaupés region are described in §2, together with the cultural setting and language attitudes. In this section I also present historical evidence and arguments in favour of shared cultural patterns.

Linguistic characteristics of the Vaupés—due to the intensive areal convergence in the context of obligatory multilingualism—are described in §3. Properties shared by the languages of the Ñana-Vaupés area are discussed in §4. The history

1 I am very grateful to Janet Barnes and Terry Malone for their comments.
of linguistic contacts and migrations in the Içana-Vaupés region is summarized in §5.

2 Linguistic Situation in the Içana-Vaupés Basin

2.1 Languages spoken

The Içana and Vaupés basins and adjacent regions, within the Upper Rio Negro basin (see map 13), contain the following genetic groups of languages:

(a) North Arawak languages spoken on the Içana and its tributary, the Aiari (Baniwa/Kurripako), on the Vaupés (Tariana) and on the Upper Rio Negro (Warekena and, formerly, Bare);
(b) several languages of the East Tucano family spoken on the Vaupés: Tucano, Tuyucu/Yururi,2 Guanano/Piratapuya, Desano/Siriano, Carapana/Tatuyo, Macana, Barasano/Taiwano, Waimaja/Bará; and one Central Tucano language, Cubeo, spoken both on the Querari river (a tributary of the Vaupés which constitutes the northern border of the Tucano area) and the Upper Aiari (see chapter 7);
(c) three Makú languages: Dàw, Hupda-Yuhup and Kakua-Nukak (see chapter 9).

There is no proof of genetic relationship between any of these language groups. The Northern branch of Arawak on the Vaupés is represented here by Tariana. Although today over 1,500 people identify as Tariana (Rodrigues 1986), the language is spoken by only about 100 people (none of them children).

Other North Arawak languages currently spoken in the Upper Rio Negro area are:

(i) dialects of Baniwa of Içana (also known as Kurripako); spoken by 3,000–4,000 people on the Içana and its tributaries and in the adjacent regions of Colombia and Venezuela (see the lists of dialects in Nimuendajú 1950/5, Rodrigues 1986).
(ii) Warekena, a dialect of Baniwa of Guainia (mainly spoken in Venezuela); spoken by a few dozen old people on the Xié river (see Aikhenvald 1998);
(iii) Old Warekena (also called Warena, or Guarequena); spoken by a few

2 Pairs of names separated by / are mutually intelligible and can be considered dialects. Terry Malone (p.c.) gives the following percentages of shared lexicon (based on the list of over 300 lexical items found in Huber and Reed 1992): Tuyucu and Yururi, 91%; Guanano and Piratapuya, 94%; Desano and Siriano, 89%; Carapana and Tatuyo, 96%.

Table 14.1 The region of Brazilian Vaupés: languages and their speakers

<table>
<thead>
<tr>
<th>Language</th>
<th>People</th>
<th>Language Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tucano</td>
<td>4,500</td>
<td>200</td>
</tr>
<tr>
<td>Piratapuya</td>
<td>1,232</td>
<td>200</td>
</tr>
<tr>
<td>Guanano</td>
<td>1,500</td>
<td>150</td>
</tr>
<tr>
<td>Desano</td>
<td>1,800</td>
<td>300</td>
</tr>
<tr>
<td>Cubeo</td>
<td>3,000</td>
<td>700</td>
</tr>
<tr>
<td>Tuyucu</td>
<td>723</td>
<td>100</td>
</tr>
<tr>
<td>Tariana</td>
<td>1,500</td>
<td></td>
</tr>
</tbody>
</table>

old people on the Xié river, and by a few hundred people in Venezuela where they moved early this century (Aikhenvald 1998);
(iv) Bare; once the most important language along the Upper Rio Negro, but now almost extinct (see Aikhenvald 1995a).3

The North Arawak languages of the Içana and Vaupés and the adjacent regions of the Upper Rio Negro basin fall into three distinct genetic groups: Warekena, Bare and Baniwa-Tariana.

East Tucano languages are closely related, but not mutually intelligible. The Tucano live along the Paca, Papuri, Vaupés and Tiquie rivers. In the Brazilian Vaupés, Tucano is gradually gaining ground as a lingua franca of the area. Table 14.1, based on Grimes (1988), gives an idea of the numbers of those who identify themselves as belonging to East Tucano groups and to Tariana, and of those who actually speak the language (see chapter 7).

Semi-nomadic Makú (the ‘untouchables’ of the region) display a number of cultural divergencies from East Tucanos and Tariana – see §2.2.

There is one Makú language, Nadëb, spoken by a few hundred people outside the Içana-Vaupés basin (see chapter 9 and Weir 1984). It does not belong to the Vaupés geographical and linguistic area. The data on Nadëb are crucial for distinguishing between areal and genetic features in the Makú languages of the Upper Rio Negro area.

Lingua francas spoken in the region are Língua Geral, or Nheengatú, and Tucano. Língua Geral is spoken in the whole region of the Upper Rio Negro (see Rodrigues 1986, Bessa Freire 1983, Moore, Facundes and Pires 1994). On the Vaupés, it is only understood by older people. Língua Geral is a creolized version of Tupinambá (Tupi-Guarani family) which was spread from the east coast of Brazil by white merchants and missionaries. It was the lingua franca of the whole Amazon region from the late seventeenth century up to the middle of the nineteenth century, and the influence of it can still be seen in a few loan words in Tariana and

3 A number of other Arawak languages of the Upper and Middle Rio Negro region became extinct during the nineteenth and twentieth centuries, e.g. Munau, Wira, Waimauna, Yurana, Passe, Yahauna, Mandawaka (Mawaca), Amarizana (Koch-Grünberg 1911), Bahiana. See map 2 and table 3.1 in chapter 3.
other languages of the Vaupés. It was gradually replaced by Tucano as a lingua franca from the early twentieth century, as a result of the language policy of Catholic missionaries and the civil authorities.

Indo-European languages are represented by Portuguese and Spanish. They are also gaining ground as lingua francas of the region, especially among younger people.

The Yanomami — semi-nomadic hunters and gatherers who live in the jungle around the minor tributaries of the Upper Rio Negro and sometimes travel as far westwards as the Vaupés region — are not culturally integrated with other peoples of the Upper Rio Negro, and do not form a part of the Içana-Vaupés linguistic area (see map 12 in chapter 13). They rarely interact with other tribes, being feared and avoided. Their status can be compared to that of Gypsies in the Balkans, who, although present in the region, do not participate in linguistic and cultural exchange.

2.2 Cultural setting and language attitudes

Multilingualism used to be — and to a great extent still is — a cultural norm over the whole Vaupés region. Among East Tucano peoples and Tariana, marriage is exogamous. There are strict marriage rules, which are governed by language affiliation (see Sorensen 1967; Aikhenvald 1996a). Marrying someone who belongs to the same language group is considered akin to incest. Jackson (1974: 62), referring to the linguistic attitudes of the Bará, reports being told: 'My brothers are those who share a language with me', and 'We don't marry our sisters.' Consequently language — which is acquired through patrilineal descent — is a badge of identity, together with the patrilineal descent. An Indian always identifies with their father's tribe and language.

The main unit, both socially and linguistically, is a speech community identified by one's father's language. Each settlement — whether longhouse or village — is multilingual because of the exogamous marriage rules. An individual generally knows between three and ten other languages of the region, including his mother's language which would often also be his wife's language, and in addition Portuguese and/or Spanish. Since language identity is a badge of ethnic identity, languages — even the closest ones — are kept strictly apart (see Sorensen 1972: 82). This creates a very strong impediment to lexical loans, in contrast to other multilingual situations, e.g. the village of Kupwar (Gumperz and Wilson 1971: 161-2), or multilingual Jewish communities (Weinreich 1964).

This unusual linguistic diversity is accompanied by a lesser degree of cultural diversity. The cultural homogeneity is corroborated by (a) shared means of subsistence, food, and ways of life; (b) shared stories and rites, and traditional social structure; with a minimal division of 'labour' between tribes (see Galvão 1979 for the differences in pottery and basket-weaving techniques between Tucano- and Arawak-speaking groups). However, each group — and apparently each subtribe — do have their own versions of origin myths.

The traditionally nomadic Makú peoples are not included in the marriage network; however, they are integrated into the economic exchange system, their main functions being as the traditional producers of arrow poison, and as skilful hunters. They have no agriculture, and depend on other peoples of the region for the staple food, manioc. They are considered inferior by East Tucano and Tariana and are denominated 'underlings' (see Martius 1994, Silverwood-Cope 1990, and chapter 9). They do not intermarry with either East Tucano or Tariana. However, they share a number of linguistic characteristics with other languages of the Vaupés region. They also share some important myths with the Tucano and the Tariana. Unlike the Tucano and the Tariana, they exhibit no inhibition against lexical loans.

As will be shown below, there is a large degree of grammatical and phonological convergence among the indigenous languages of the Vaupés. The impact of these convergence phenomena is especially instructive in the case of Tariana, since they are discernible as the result of 'foreign' influence.

The main difference between the Içana basin and the Vaupés basin is the lack of compulsory linguistic exogamy in the Içana. Baniwa cohabit with Central-Tucano-speaking Cubeo on the Aitari river, a tributary of the Içana, and on the Querari, a tributary of the Vaupés which borders the Içana basin. The Cubeo and the Baniwa are in constant contact which results in diffusion of cultural and linguistic features. Marriage tends to be exogamous between different descent groups and dialectal groups of the Cubeo and the Baniwa. The Cubeo have assimilated Baniwa dances and chants (see Goldman 1979, Gomez-Imbert 1996). Some of the Cubeo-speaking groups may originally have been Baniwa. Nowadays there are no — or almost no — regular contacts between the East Tucano of the Vaupés and the Tariana, on the one hand, and the Baniwa on the other.

4 In most cases a speech community is identical to a phratry, the members of which can intermarry. However, 'the Cubeo, Piratapuya and Macuna are exceptions, in that each language represents more than one phratry, so that, for example, Cubeos can marry Cubeos if the marriage partner is a member of an exogamous phratry' (Malone 1987: 27). See also Grimes (1985). Each phratry is divided into endogamous sibs. There are minor linguistic differences between sibs (Terry Malone, Janet Barnes, p.c.).

3 Within a larger language group, all phratries are ranked, and there is a ranking of sibs within a phratry. There is a tendency for lower-ranked sibs to be those whom one would suspect were originally speakers of non-Tucano languages, in most cases Makú (see §1 of chapter 9 in this volume; Janet Barnes and Terry Malone, p.c.).
However, the Baniwa share a number of cultural similarities with the indigenous groups of the Vaupés which may be indicative of older contacts and of diffusion. These similarities include the structure of the kinship system (described in Oliveira 1975, Galvão 1979), patrilineal descent, and the hierarchical organization of descent groups. There are also a number of traditions and myths shared by the Baniwa and the Tucano, e.g. the myth of a woman going to the ‘beyond’ with the help of an agouti (C. Hugh-Jones 1979); stories about the cunning turtle and the tapir; the story of the evil spirit’s son who turned into a deer; the cult of the magic flute, Jurupary (S. O. Hugh-Jones 1979); burial of the dead inside the house; infanticide of twins; and the festival of Dabukuri – an offering accompanied by drinking manioc beer and ritual dances (cf. Reichel-Dolmatoff 1986). There are also a number of linguistic features shared by Baniwa, Tariana and East Tucano languages but not found in other North Arawak languages. These will be considered in §2.4.

2.3 Historical evidence

Historical and traditional evidence shows that neither the Tucano nor the Tariana constitute the autochthonous population of the Vaupés.

According to Nimundajú (1982), the original inhabitants of the Vaupés area were Makú tribes, and East Tucano tribes then entered the area from the west. The Tariana are the most recent arrivals. They probably came into the Vaupés basin from a tributary of the Ícana river – probably the Aiari (around the end of the sixteenth century according to Brúzzi 1977, Koch-Grünberg 1911 and Nimundajú 1982). This description of the origin of the Tariana is reflected in their myths, according to which all of them originated from a ‘hole’ at the waterfall of Apuí, on the Aiari river, where at the time they lived together with the Baniwa, and with the Cubeo.

Numerous stories about the wars between the Tariana and the Guanano, and between the Tariana and the Desano, provide rich ethnographical evidence for the Tariana invasion of the Vaupés (see Brúzzi 1977). The Desano, as the inhabitants of the main rivers, were among the first to have suffered from the invasion of the Tariana. I was told by my teachers of Tariana that, as the result of the Desanos’ defeat, they were designated the ‘younger brothers’ of the Tariana, and this is why the Tariana do not marry them.

We can reconstruct the following historical scenario for the linguistic situation in the Vaupés during the past 500 years.

(i) Before 1500: East Tucano tribes moved from the west into the Vaupés area, which was previously inhabited by Makú tribes (cf. chapter 9). The East Tucano established dominance over the Makú. The spread of Tucano-speaking peoples to the Ícana basin may have occurred at about the same time. According to the oral tradition of the Cubeo (Gomez-Imbert 1996), they came to the Querari from the Ícana tributaries a few hundred years ago, chased by Baniwa-speaking tribes. Some of the Cubeo speakers on the Querari are descendants of Arawak-speaking tribes. It is hard to establish the exact age of Baniwa–Cubeo contacts in the Ícana basin; however, it was sufficiently long ago to promote the areal diffusion discussed below.

The Baniwa of Ícana were at this time living together with the Tariana in the Ícana basin.

(ii) Around 1600: the Tariana moved from the tributaries of the Ícana river – thus splitting from the Baniwa of Ícana – to the Vaupés region, in which Tucano tribes were already established. This marked the beginning of contact between East Tucano and Tariana.

(iii) Around 1750–80: the first contacts with Portuguese took place, which started the spread of Língua Geral as a lingua franca.

(iv) Around 1900: Tucano started to gain ground as a lingua franca of the area, with some Tariana settlements beginning to use mainly Tucano (see Koch-Grünberg 1911). This tendency increased with the establishment of permanent Salesian missions in the Vaupés in 1925, and resulted in the growing endangerment of indigenous languages other than Tucano in the Vaupés region, and the growing obsolescence of Língua Geral.

The main consequence of the spread of Tucano in the Brazilian Vaupés is the gradual undermining of the identification between language and tribe. Language has gradually ceased to be an emblem of tribal identity, and the majority of languages other than Tucano have become endangered. The discrepancy between the number of those who belong to a tribe and those who actually speak the language (see table 14.1) is particularly marked in the case of Tariana. The spread of Tucano is also leading to the gradual disappearance of one of the most fascinating multilingual areas of the world, and the areal phenomena associated with it.

This does not seem to be the case in the Colombian Vaupés. Tucano, however, is also the prestige language there; when the Tucano are present, speakers of other languages use Tucano. The language which seems to be disrupting the Tucano sociolinguistic situation there is Spanish (Ardila 1989; Terry Malone p.c.).

3 THE VAUPÉS REGION AS A LINGUISTIC AREA

3.1 General observations

Here we consider the properties of languages spoken in the Vaupés region which characterize it as a linguistic area.
As pointed out by Thomason and Kaufman (1988: 96ff.), a long-term multilateral Sprachbund seems to promote gradual isomorphism in all aspects of the structure of languages except for the phonological shape of morphemes. A famous example of the same surface structure for several languages spoken in the same linguistic area comes from Indo-Aryan and Dravidian languages of Kupwar (Gumperz and Wilson 1971). In the case of numerous multilateral linguistic areas, such as the Balkans (Thomason and Kaufman 1988: 97), the direction of diffusion is easily established. In the case of the Vaupés, it can also be established for certain features. The following characteristics of the Vaupés as a linguistic area should be borne in mind:

- the fact that there are several East Tucano languages and only one Arawak language has provoked, by and large, unilateral diffusion patterns: from East Tucano into the Arawak language, Tariana;¹
- there is a recent tendency of Tucano to develop domination over other languages, hence the destruction of multilingualism in the community;
- language identification and the perceived link between tribe and language drastically limit the extent of lexical borrowing;
- while considering areal diffusion patterns and structural convergence between the languages of the area, one has to take into consideration language obsolescence phenomena for some languages.

The Vaupés linguistic area (for some useful comments see Brüzzi 1967, 1977) has multilateral multilingualism characterized by linguistic diversity and a comparative lack of cultural diversity. As already mentioned, the languages belong to three genetically unrelated families – North Arawak, Tucano and Makú – and their speakers do not accept language-mixing. However, people tend to identify the language with its forms and the prohibition on ‘language mixing’ only relates to forms, not to grammatical categories or structural patterns. As a result there is a great degree of diffusion of structural patterns but little borrowing of actual forms.

The idea that the Vaupés region constitutes a linguistic area was first suggested by Sorensen (1967), with reference to the Colombian side, where only East Tucano languages are spoken. Sorensen pointed out the existence of a bundle of morphosyntactical isoglosses which constitute an East Tucano ‘profile’ of language (1972: §2-3). However, whether this East Tucano profile is due to areal diffusion patterns or to the common genetic origin of East Tucano languages remains a problem which goes beyond the scope of the present discussion. (The solution of this problem would involve a full reconstruction of proto-East-Tucano and comparison of it with proto-West-Tucano and proto-Tucano.)

The important difference between the Brazilian Vaupés region and the Colombian Vaupés region from the point of view of areal diffusion phenomena is the fact that a non-East-Tucano language, Tariana, is still spoken in Brazil. Thus, the East-Tucano-like features in Tariana (and also, North-Arawak-like features in East Tucano) enable us to establish a direction of diffusion in the area.

For the analysis I will use the following kinds of data:

- my own field data on Tariana and on other North Arawak languages of the wider region (Baniwa of Íçana, Warekena of Xié, Bari); proto-Arawak reconstruction and internal reconstruction (see Payne 1991; Aikhenvald 1994a, b, 1995a, 1996a, b, in prep.);
- a number of descriptions of East Tucano languages (not all of equal quality), and my own field data, and also some data on proto-Tucano reconstruction and descriptions of East Tucano and West Tucano languages which do not participate in the Brazilian Vaupés as a linguistic area, as well as the data on proto-West-Tucano.

The analysis of areal and inherited features of Tariana includes a fair amount of internal reconstruction, and synchronic comparison with Baniwa of Íçana and East Tucano languages. A comparison of Tariana with geographically close and genetically related North Arawak languages is extremely useful, especially if we are able to determine which properties Tariana shares with East Tucano languages, and which are shared by other North Arawak languages, but absent from Tariana (also see Aikhenvald 1996a, on innovations and language attrition phenomena in Tariana).

Contact between Tucano and Makú languages (Dâw, Hupda, Yuhup) in the Vaupés appears to be much older than contact between Tucano and Tariana – see §2.3. Note that one would expect mainly unilateral diffusion patterns from East Tucano to Makú, since a fair proportion of Makú know at least one East Tucano language, or more, but East Tucano peoples seldom have competence in a Makú language.

To establish areal diffusion patterns in the Makú languages spoken in the Vaupés.
and Upper Río Negro regions, data from Nadêb, a Makú language spoken in the Middle Río Negro area (based on Martins 1994; Weir 1984, 1990; see also chapter 9 above) are of crucial importance.

The materials available on Makú languages spoken in the Vaupés and Upper Río Negro areas are not extensive. However, they reveal that, in most cases, areal diffusion is more superficial (compared with the Tariana–Tucano interaction). This is what would be expected, since the Makú are accorded an inferior social status and are not fully integrated into the multilingual socio-cultural community. Further work is needed to determine possible areal diffusion features in the Makú languages spoken in Colombia (Kakua and Nukak).

I shall now consider shared phonological characteristics of the languages of the Vaupés (§3.2), their grammatical structure (§3.3), syntax and discourse techniques (§3.4) and semantics (§3.5).

3.2 Phonological characteristics

A number of phonological characteristics of Tariana are shared with East Tucano languages, but not with the genetically related North Arawak languages from outside the Vaupés area. Some phonological characteristics of Makú languages are similar to East Tucano languages, but differ from Nadêb, another language of the same family. They can be explained through areal diffusion in the Tucano languages. The following characteristics are shared by Tucano, Tariana and Makú languages.

1 Nasalization is a word-prosodic feature in Tariana and in Hupda-Yuhup and Kakua (see chapter 9). The same phenomenon is found in all East Tucano languages, but not in other North Arawak languages or in Nadêb (see Aikhenvald 1996a). The origin of the direction of nasal prosody (from left to right or from right to left) remains problematic. Left to right is the pattern attested everywhere; however, Terry Malone (p.c.) reports that some traces of the opposite direction (right to left) can be found in Desano, Siriano, Barasano, and also in one West Tucano language, Koreguaje. Note that in Choco languages, spoken in the Andes, nasalization operates from left to right; this may be indicative of areal diffusion between Choco and Tucano (Malone p.c.).

2 Pitch accent and intonation patterns in Tariana and in the Makú languages display striking similarities to the ‘Tucano’ accent in Tariana. Dâw and Hupda-Yuhup have two tones, and Kakua has four (see chapter 9). Tones are absent from Nadêb. Pitch accent is attested in Baniwa (see §4); however, unlike the languages of the Vaupés there is here no tendency towards developing a tone distinction.

3 There are two phonemes with low functional load in Tariana which are also attested in Tucano languages and in all Makú languages, but are absent from other North Arawak languages. The vowel i appears in two morphemes: -pi ‘augmentative’ and -ihmeni ‘to moan’; and o appears in the feminine marker on kinship terms: -Co (cf. Baniwa -Cu, Tucano -Co) (where C stands for a consonant).

The following features of Tariana result from East Tucano influence.

A In Tariana, inherited phonemes absent from Tucano – aspirated stops and aspirated glide – have a reduced functional load with respect to other phonemes and when compared to the status of the same phonemes in other North Arawak languages (cf. Aikhenvald 1995a for Bare; Taylor 1990, Aikhenvald 1996b, for Baniwa). They also tend to be in free variation with non-aspirated stops and glide respectively. There is a tendency to use just s instead of ʃ and tf; s and tf occur in all North Arawak languages, and Tucano languages have only s. Differentiation in vowel length, which is absent from Tucano, also has a low functional load in Tariana. There is significant variation between long and short vowels in Tariana, e.g., Tariana wuni-uní ‘water’, Baniwa wuni. As is typical for a language contact situation, Tariana is losing the phonemes not found in its neighbours.

B Syllable patterns in Tariana are: (C)V, (C)Vh; the latter alternates with (C)hV. Unlike in other North Arawak languages, h in an unstressed word-initial or morpheme-initial syllable can occupy a postvocalic position, thus creating VC-like syllables, e.g., ekùkùpi–ekúkúpi ‘day’ (cf. Baniwa hekúpi ‘day’). There are several cases where a syllable CVh occurs word-medially (independently of stress pattern) in slow to normal register. The onset of the following syllable must be a voiceless stop ʃ or k, e.g. -nihtá ‘to think, to reason’, -buhtá ‘conditional’, marátahka ‘a kind of wave’, karáhta ‘lung’. In rapid speech ʃ disappears; maráta ‘a kind of wave’, -buhtá ‘conditional’, etc. All these words have a North Arawak origin. The occurrence of CVh syllables in Tariana resembles CVh syllables in Tucano. Note that in East Tucano languages CVh and/or CV? are the only instances of consonantal coda.

C Tariana has several phonological characteristics not found elsewhere in North Arawak but are characteristic of Tucano:

- y→ɪ contiguous to nasal vowels.

In word-initial and word-medial position Baniwa dz and Kurripako y correspond to Tariana ì, if the following or the preceding syllable contains a nasal consonant, e.g., Tariana ìma, Baniwa dzma.

Only Guanano has been analysed as having phonemic aspirated stops (Waltz and Waltz 1967). In other Tucano languages, there is a tendency to slightly aspirate voiceless velar stops before non-front high vowels (Terry Malone p.c.).
Kurripako *yama* 'two'; Tariana -*šami*, Baniwa -*džami*, Kurripako -*yaní* 'to die, to be terminally ill'; Tariana *pamíthí*, Baniwa *pamudzu*, Kurripako *panuwa* 'middle' (cf. also Tariana *šamu*, Warekena of Xis *yama-ducevil spirit*).

Tariana *y* corresponds to Baniwa *y* in other positions, e.g., Tariana, Baniwa *malájí* 'knife'. Tariana *y* corresponds to Baniwa, Kurripako *y* in loan words from Língua Geral, e.g. Tariana, Baniwa *yamaná* 'a white person'; Tariana, Baniwa *yáxíri* 'caxiri, manioc wine'.

- *y>* word-initially, e.g., Tariana *yuru* 'a kind of mosquito' realized as [dyururj] in rapid speech.
- *d>* contiguous to a front vowel (optional), e.g., Tariana -*pídena*-*pírena* 'remote past inferred'.

The fact that these phonological processes are found in other South American languages, as well as in languages of other areas of the world, does not go against their areal character in Tariana, since they are absent from other North Arawak languages; the same holds true for nasalization as a word-prosodic property, and for glottal stop.

D. The glottal stop, absent from other North Arawak languages, sporadically appears in Tariana word-finally (but is not contrastive). There is a glottal stop phoneme in Desano, Siriana, Tucano, Guanano and Piratapuya, but in Barasano and Tatuyo the glottal stop occurs as a sentence-final prosodic feature (not as a phoneme) (Jones and Jones 1991; Whisler and Whisler 1976).

### 3.3 Grammatical structure

#### 3.3.1 Typological profile and word structure

East Tucano languages are head-marking, with elements of dependent-marking. They are agglutinating and suffixed.

Arawak languages are head-marking, and agglutinating with elements of fusion. They are suffixing with the addition of one prefix position in a word. North Arawak languages tend to have few suffix positions in verbs, Tariana is basically head-marking, with a few elements of dependent-marking, one prefix position, and a large number of suffix positions in verbs.

The Makú languages of the Vaupés combine elements of head-marking and dependent-marking. They are predominantly suffixing, with very few relics of a prefixing morphology. In contrast, Nadeb is predominantly head marking, with an unusually large number of prefix positions for an Amazonian language. The almost complete loss of prefixing morphology in Makú languages other than Nadeb can be considered the result of areal diffusion from Tucano, with its predominantly suffixing structure (see chapter 9 on Makú).

---

### 3.3.2 Nominal morphology

#### (i) Case-marking

Core-case marking on personal pronouns with an animate referent – and on nouns, dependent on their topicality – is the most striking areal property shared by East Tucano languages, Tariana and Makú in the Vaupés.

Generally, Arawak languages do not have case-marking. Tariana has the following case-marking system which shows similarities to East Tucano languages (see Aikhenvald 1994b; on case-marking in Tucano languages, see Kinch 1977, Gralow 1980).

- Personal pronouns with an animate referent have an obligatory opposition between subject case (used to mark A/S) and non-subject case, used to mark O and dative argument.
- Other constituents (including third person pronouns with inanimate referent) have an optional topic marker (Tariana *-nuku*, East Tucano *-re*) which goes on any non-subject constituent provided it is the topic.

In East Tucano, case marking is obligatory with personal pronouns and highly desirable with personal names. Note that the non-subject topic marker is optional in (1) but obligatory in (2)-(4).

**Piraputanga** (*my fieldwork*)

(1) *yey oha-gi-(re) Lenita-re*  
1sg write-RELATIVE-(TOPICAL NON.A/S) Lenita-TOPICAL NON.A/S ou-’u  
give-PAST  
'I gave the pen to Lenita.'

(2) *yey bìFF-re Lenita-re ou-’u*  
1sg YOU-TOPICAL NON.A/S Lenita-TOPICAL NON.A/S give-PAST  
'I gave it to Lenita' (talking to a baby)

**Tariana**

(3) *nuha pa-dana-ni-na-nuku*  
1sg IMPERS WRITE-TOPIC ADVANCING-CL:VERTICAL TOPICAL NON.A/S Lenita-nuku nu-a  
Lenita-TOPICAL NON.A/S 1sg-give  
'I gave a pen (lit. a vertical writing device) to Lenita.'

---

9 Note that in chapter 7 the case marking system in Tucano is described as a system of 'specifiers' relating to the correlation between the choice of case and the discourse function of a noun.
The languages of the Vaupés region tend to have one peripheral case-marker or postposition. Most East Tucano languages have one peripheral case-marker -pi (Bersano -bi) 'locative/directional, clative' which can co-occur with the definite/topical non-subject marker -re if the corresponding constituent is topical. Tariana has a corresponding suffix -se, e.g., Tucano Bogota-pi-re, Tariana Bogota-se-nuku (Bogotá-LOC-TOPICAL.NON.A/S) 'to/from in Bogota'. Daw, Hupda-Yuhup and Kakua have just one locative marker (D -rid, Kk -p 'to, from, in') (see chapter 9). Nadèb, which does not belong to the area, distinguishes ablative and locative (Weir 1984: 96).

The following properties are shared by Tariana and Tucano languages, but not by Baniwa or Makú.

(iii) **Alienable and inalienable possession**

In most North Arawak languages inalienably possessed nouns include all kinship nouns and body parts and a few other items, e.g. 'home', together with deverbal nominalizations. Inalienable possession is marked with the help of possessive prefixes, as in (7). In Arawak languages other than Tariana, alienably possessed nouns have a special possessive suffix, as in (13) (see Aikhenvald 1994b).

In Tariana, inalienably possessed nouns are restricted to body parts and just a few kinship nouns (a smaller set than in other North Arawak languages). Inalienable possession is marked by possessive prefixes. Alienable possession is shown simply by juxtaposition.

**Tariana**

(7) nu-hado or nuha ha-dua
   1sg-mother or 1sg parent-FEM
   'my mother'

Compare this with the same lexical item as inalienably possessed in Baniwa:

(8) hnadua (<underlying) nu-ha-duá
   1sg-parent-FEM
   'my mother'

In Tucano inalienable possession is marked by juxtaposition, e.g.: 

(9) yií pa-co
   1sg parent-FEM
   'my mother'

All Tucano languages use the genitive marker ya plus a classifier to mark alienable possession, e.g. Tucano:

(10) miií ya wií
    2sg POSSESS FOOD
    'your house'
Like Tucano, Tariana uses the possessive marker *ya* with cross-referencing prefixes plus a classifier as an alternative possessive construction for alienably and inalienably possessed items; the head noun is then usually omitted. In Baniwa classifiers are not used in possessive constructions (unless possession is predicative – see §4.3).

**Tariana**

(11) pi-ya-ku (ama-ku)  
2sg-POSSV-CL:EXTENDED (hammock-CL:EXTENDED)  
'your hammock'

**Tucano**

(12) mi’i ya-gi (pū-gi)  
2sg POSSV-CL:LARGE (hammock-CL:LARGE)  
'your hammock'

**Baniwa**

(13) nu-pieta-ni  
1sg-hammock-POSSV  
'my hammock'

In both Nadëb (Weir 1984: 84–6) and Dâw (Martins 1994: 40, 140), alienable and inalienable possession are marked with juxtaposition of nouns:

**Dâw** (Martins 1994: 46)

(14) dâw put  
person hair  
'person's hair'

For predicative possession, a possessive marker which combines with classifiers is used in Dâw, but not in Nadëb; this property is shared with Baniwa and will be considered in §4.3.

(iv) Gender

In other North Arawak languages, a gender opposition (feminine animate vs the rest) is found in verbal cross-referencing markers and demonstratives. Classifiers are not used with demonstratives. Tariana and Tucano have a gender opposition in verbal cross-referencing markers, personal pronouns and kinship terms. However, demonstratives have special animate and inanimate forms and they also combine with classifiers. The head noun is frequently omitted.

This is illustrated for Tariana in (15) and for Tucano in (16). In Baniwa, shown in (17), the demonstrative takes a non-feminine form, but no classifier.

**Tariana**

(15) ha-ku (ama-ku)  
DEM/NAN-CL:EXTENDED (hammock-CL:EXTENDED)  
'this hammock'

**Tucano**

(16) ati-gi (pū-gi)  
DEM-CL:LARGE (hammock-CL:LARGE)  
'this hammock'

**Baniwa**

(17) hliehê pieta  
DEM+N.FEM hammock  
'this hammock'

(iv) Classifiers and repeaters

Baniwa, Tariana and Tucano languages have noun classes, and numeral and possessive classifiers (see Aikhenvald 1994b; Barnes 1990, for Tuyuca). Baniwa and Tariana are the only North Arawak languages which use classifiers in possessive constructions. However, there are a few significant differences between Baniwa, on the one hand, and Tariana and Tucano, on the other.

In Tariana and Tucano there is a special classifier for animate nouns; in Baniwa animate non-human nouns are classified according to their shape:

**Tariana**

(18) apí/a:pi hanu-ite  
snake big-CL:ANIM  
'a big snake'

**Tucano**

(19) pinõ phai-gi  
snake big-CL:ANIM  
'a big snake'

**Baniwa**

(20) a:pi maka-khay  
snake big-CL:CURVILINEAR  
'a big snake'

12 Among the Makú languages, Kakua distinguishes masculine and feminine genders in third person singular pronominal prefixes, independent pronouns and demonstratives, probably due to the Tucano influence (see §5 of chapter 9).

13 In Tucano there is also gender marking on adjectives, demonstratives, numbers and nominalized verbs.
Baniwa has a closed, albeit large set of forty-four classifiers, including a residual classifier for unclassifiable items (see Aikhenvald 1996b). Tariana and Tucano languages have very large sets of classifiers which can be considered 'open' since just about any noun with an inanimate referent can be used to mark agreement to classify the same noun (this phenomenon is known as 'repeaters' – see Aikhenvald forthcoming: chapter 10). Tariana and Tucano use 'repeaters' for otherwise unclassifiable nouns.

In East Tucano and Tariana, 'repeaters' (i.e. classifiers which look segmentally identical to the noun) lose their stress, whereas the corresponding independent noun will always retain its stress (Barnes 1996 for Tuyuca; Ramirez 1997 for Tucano).

**Tariana**

(21) panisi hanú-panisi
house big-CL:HOUSE
'a big house'

**Tucano**

(22) wi'í phai-ri-wi'í
house big-SG.MASC-CL:HOUSE
'a big house'

**Baniwa**

(23) pānti maka-dájì
house big-CL:ROUND
'a big (round) house, a big burrow'

'Repeaters' can be used as 'ad hoc' classifiers in Tariana and Tucano (see Aikhenvald 1994a, for discussion). The use of a 'repeater' in Tariana is illustrated in (24). *Panisi* 'house' is used as an agreement marker on the adjective *matja* 'good', as an indicator of the special importance of the newly introduced referent 'house'. Otherwise, a classifier -*dapana* 'habitat' would be used, as in (25).

**Tariana**

(24) nuhua matja-panisi-mha panisi nu-na
1 good-house-PRES.NON.VISUAL house 1sg-want
'I want a really good house' [the man said to the magic ring].

(25) kayu diha di-sata-ka diha depita
so he 3sg.n.fem-ask-sEQ he night+ADV
hiku-pidana panisi matja-dapana thuya
appear-REM.PAST.INFERRED house good-CL:HABITAT all

Classifiers are widely used in anaphoric and discourse-backgrounding functions. The head noun is often omitted in Tucano and Tariana (it is omitted from just before the second occurrence of *hiku* in (25)). This does not happen in Baniwa.

However, classifier systems in Baniwa and Tucano–Tariana still display a number of shared properties not found in other neighbouring North Arawak languages. We shall return to these in §4.4.

Tariana and Tucano have portmanteau plural + classifier morphemes, absent from North Arawak.

**(vi) Plural and plural agreement**

Tariana – unlike other North Arawak languages – and Tucano distinguish between animate and inanimate plural. Plural agreement is obligatory in a noun phrase. In both Tariana and East Tucano languages, inanimate plural morphemes are not used for less than four objects. There are many suppletive plural forms, especially for kinship nouns.

3.3.3 Grammatical relations

The marking of grammatical relations is uniform among the languages of the Vaupés.

Tucano languages have straightforward nominative–accusative patterns in cross-referencing and case-marking (see §2 of chapter 1 above).

Tariana has lost the proto-Arawak morphological split-ergativity, which is marked with cross-referencing affixes in Baniwa of Ícana and Warekena (for further discussion of cross-referencing affixes and enclitics in Baniwa of Ícana, Warekena, Bare and Tariana see Aikhenvald 1995b). In these languages:

- prefixes cross-reference A (in a transitive clause) and S (in an active intransitive);
- suffixes and enclitics cross-reference O in a transitive clause and S (in a stative intransitive).
Tariana has no cross-referencing enclitics, but it has acquired a nominative-accusative case-marking pattern under areal pressure from Tucano.

One of the striking differences between Nadëb and the Makú languages of the Vaupés (Daw and Hupda) lies in the marking of grammatical relations. Nadëb has one of the most complicated ergative-type systems in the world (Dixon 1994: 134, 178, 233; and chapter 9 above). Daw and Hupda are consistently nominative-accusative (see §2 of chapter 9).

This may be accounted for by areal pressure from the predominantly suffixing East Tucano languages. Unlike Nadëb, Daw and Hupda have strong verb-final tendencies. East Tucano languages and Tariana are also predominantly verb-final. This drift towards AOV/ SV typology due to areal diffusion may have led to the loss of constituent order being used to mark grammatical relations (as it does in Nadëb). Daw and Hupda have acquired accusative patterns of case-marking which are similar to East Tucano, and to Tariana.

3.3.4 Verbal morphology and predicate structure

(i) Verb serialization

Among the languages of the Vaupés basin, Tariana and Tucano have a typologically unusual combination of verb serialization and verb compounding. Serial verb constructions can consist of a sequence of two or more verbs. Verb compounding results in the creation of aspect- and valency-changing markers (Aikhenvald ms. Gomez-Imbert 1988).

The Makú languages of the Upper Rio Negro have productive verb compounding which results in the creation of various aspect markers. This property is shared with East Tucano and Tariana, but not with the genetically related Nadëb. It is also the means whereby Makú languages are gaining 'new' morphology to make up for the prefixal morphology they lost under areal pressure from East Tucano suffixing-type languages.

The following features are shared just by East Tucano and Tariana.

(1) Unlike other North Arawak languages, Tariana is like East Tucano in distinguishing eyewitness and non-eyewitness evidentiality, immediate and distant future, recent and remote past, conditional mood and several imperatives.
(2) Tucano and Tariana have suffixal negation, a negative copula, and special negative words 'I do not know', 'there is nothing'.
(3) Tucano languages have a number of compound verbal forms which consist of a participial form of the main verb and an auxiliary. In relaxed speech by younger people in Tariana, serial verb constructions are replaced by a sequence 'relative form of the main verb plus auxiliary-like stance verb', following the Tucano pattern of a 'participial form of the main verb plus auxiliary' (Sorensen 1972).

3.4 Syntax and discourse techniques

The most striking feature of syntactic organization of the languages of the Vaupés is a strong verb-final tendency. This is found in East Tucano, Tariana and in Makú. Other North Arawak languages show a verb-medial tendency; Nadëb is predominantly verb-initial (with constituent order used to mark grammatical relations: see §10 of chapter 9).

The following properties are shared by Tariana and Tucano, but not by Makú.

(1) Tariana and Tucano have long chains of juxtaposed clauses, and use relative verbal forms ('participles') to mark the predicate of a subordinate clause. Other North Arawak languages prefer sequencing verbal clitics (similar to medial verbs in Papuan languages; cf. Aikhenvald 1995a, 1998, on sequencing clitics in Bare and Warekena). Tariana still makes limited use of a sequential clitic -ka, to mark the predicate of a complement clause (and this is used most often in narratives told by old speakers). Thus, Tariana combines both 'Tucano-like' and 'North-Arawak like' strategies of complementation.

(2) Tariana – unlike other North Arawak languages – and Tucano have a complex system of switch-reference. This is illustrated in (26) and (27) from Tariana. In (26), -nisawa is used to mark different subjects, and in (27) -sita marks the same subject.

(26) kay-di-ni-niswa
so-3sg.n.fem-do-AFTER-DS
he shore-LOC
3sg.n.fem-enter
di-a-pidana
3sg.n.fem-go-REM.PAST.INFERRED
3sg.n.fem-enter
3sg.n.fem-go
'After he (the man) did this, he (the otter) came up on the shore and entered (water) and dived away.'

(27) duhua du-mara-pidana
du-mara-sita
she 3sg.fem-float-REM.PAST.INFERRED 3sg.fem-come-AFTER-SS
du-nu-ka
duma
3sg.fem-come-DECL
3sg.fem +look for
dhewi-nuku
3sg.n.fem + pip-TOPICAL.NON.AS
'After she floated downstream, she came to look for the pip (of a peachpalm fruit).'

14 A subordinating enclitic -ka is also found in four East Tucano languages (Carapana, Tatuyo, Waimaja and Yururi), it is not clear whether or not it was spread by contact.
4 Other North Arawak languages – like most other South American Indian languages – avoid sentences with two full NPs, especially when one of them is a free pronoun. Free pronouns are mainly restricted to emphatic function. Tariana, similarly to East Tucano languages, makes wide use of personal pronouns, as in (26) and (27).

4.1 Pitch accent

Vaupés.

15 languages properties. Only one linguistic properties shared by Baniwa besides cultural similarities widespread across this area, there are a number of linguistic properties shared by Baniwa and the languages of the Vaupés discussed above, but absent from other North Arawak languages. Table 14.2 summarizes these properties. Only one of these properties - pitch accent - is also shared by the Makú languages of the region.

3.5 Semantics

Tariana has undergone a number of lexical shifts under East Tucano influence. In Tariana, one word - *keri* - is used to refer to both 'sun' and 'moon', similar to *mshigpun* Tucano. Other North Arawak languages have two distinct words for these, e.g. Baniwa of *Içana kegi* 'moon', *komu* 'sun'. Tariana *iri* 'blood' has acquired the meaning of *'sap*' (hence: rubber, plastic) under the influence of East Tucano *di*. Other North Arawak languages have different words for 'blood' and *'sap*', e.g. Baniwa of *Içana iri* 'blood', *dzeka* 'sap, rubber, plastic' (see Aikhenvald 1996a). Dáw and Hupda-Yuhup also use just one term for 'moon, sun', whereas Nédoeb has two different words (see chapter 9 on Makú).

4 PROPERTIES SHARED BY LANGUAGES OF THE IÇANA AND VAUPÉS REGION

We can now look at the larger area consisting of the Içana and the Vaupés basins. Besides cultural similarities widespread across this area, there are a number of linguistic properties shared by Baniwa and the languages of the Vaupés discussed above, but absent from other North Arawak languages. Table 14.2 summarizes these properties. Only one of these properties – pitch accent – is also shared by the Makú languages of the region.

4.2 Topic-advancing verbal derivation

Tariana and Baniwa have an argument-manipulating derivation which creates an alternative construction type. There is no difference in transitivity value or grammatical relations; it is just that a different argument is focused on in the basic and alternative construction types (Dixon and Aikhenvald 1997: 91-4). This derivation, marked with suffix *-ni* on the verb, has the following properties.

(i) The *ni*-derivation marks the advancement of any non-subject topical constituent to subject; this constituent cannot take a non-subject case marker (*-na* for pronouns; *-nuku* for everything else).

(ii) The original *A/S* is always cross-referenced on the predicate; the full NP is often omitted.

(iii) The *ni*-derivation can be formed on intransitive active (*S*<sub>1</sub>) verbs and transitive verbs (transitivity is maintained), but not on stative prefixless (*S*<sub>2</sub>) verbs.

Table 14.2 Properties shared by languages of the Içana-Vaupés

<table>
<thead>
<tr>
<th>Properties shared</th>
<th>Vaupés</th>
<th>Içana</th>
<th>North Arawak languages outside this area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch accent</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Topic-advancing verbal derivation</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Possessive classifiers</td>
<td>yes yes</td>
<td>no</td>
<td>yes (with possessive predicates)</td>
</tr>
<tr>
<td>Possessive -ya- to which classifiers are attached</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Classifiers with demonstratives</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Several types of classifiers</td>
<td>yes yes</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

absent from Warekena and Bare, North Arawak languages spoken outside the area (Aikhenvald 1995a, 1998).
The verb includes a classifier suffix which cross-references just the NP moved into surface subject slot.

Other argument NPs (besides the original subject and the NP promoted into the surface subject slot) are maintained.

The main discourse function of topic advancement is to mark a constituent which is more topical than the underlying subject.

Example (28) is a plain transitive clause; the verb is underlined.

Example (29) shows a direct object.

The advanced constituent has five of the six criterial properties of 'subject': (i) equi-NP deletion; (ii) subject position in constituent order: AOV/SV; (iii) pronominalization; (iv) agreement in classifier; (v) case-marking. Cross-referencing is the only subject property retained by the original subject, e.g. 1sg in (29).

Tucano has a similar 'focus-changing' derivation marked with a suffix -nu' (West 1980: 59; Terry Malone p.c.16). The suffix precedes the negative -ni and tense-aspect morphemes. This derivation applies to transitive and to intransitive verbs. In East Tucano languages grammatical subjects are expected to be topical; in the rare cases where a non-subject (most often, the object) is more topical than the subject, the -no derivation must be used (Terry Malone p.c.). The constituent advanced to subject by a -no derivation triggers subject agreement in gender/number on the verb, and cannot be marked by the case morpheme -re 'definite/topical non-subject'.

14 Içana-Vaupés linguistic area

Example (31) illustrates an intransitive clause; in (32) the -no derivation has been applied, with a locative constituent being advanced to subject.

(iii) equi-NP deletion; (ii) subject position in constituent order: AOV/SV; (iii) pronominalization; (iv) agreement in classifier; (v) case-marking. Cross-referencing.

The following examples illustrate -no with a transitive verb. Example (33) is a plain transitive clause. The direct object is more topical than the subject in (34), and then the -no derivation has to be used.

Example (35) illustrates the use of a plain transitive clause.

4.3 Possessive classifiers and -ya- possessive marker

Tucano and Tariana use possessive marker -ya- with classifiers to form possessive constructions (§3.3). In Tucano, these are restricted to alienable possession (see (10) above). Possessive constructions marked with -ya- are found in Tariana, where they can be used with both alienably and inalienably possessed nouns. Example (11) above shows an alienably possessed noun; (36) exemplifies an inalienably possessed one:

Similar derivations are found in other East Tucano languages (Terry Malone p.c.).
Baniwa also has possessive marker -yal-dza⁰ which is used exclusively to mark predicative possession of alienably and inalienably possessed nouns. Verbal classifiers are obligatory in this construction.

(37)  huiehe  pieta  nu-dza-ku
     this:SG.N.FEM  hammock  1sg-POS-CL
   'This hammock is mine.' (alienable)

(38)  huiehe  hnuui-da  nu-dza-da
     this:SG.N.FEM  head-CL:ROUND  1sg-POS-CL:ROUND
   'This head is mine.' (inalienable)

This construction is not found in any other North Arawak language. It is found elsewhere in Tucano. However, in Baniwa it is restricted to possessive predication. This suggests the possibility of unilateral diffusion from Tucano into Baniwa and Tariana within the Vaupés-Içana contact area.¹⁰

4.4  Complex systems of classifiers

East Tucano languages, and Baniwa and Tariana, are unusual in having extensive systems of classifiers in which the same or almost the same sets of classifier morphemes are used in a number of distinct morphosyntactic environments. All of these languages use classifiers as agreement markers on adjectives and numerals. The difference between the possessive classifiers in Baniwa and those in Tucano–Tariana was described in §4.3. Unlike Baniwa, Tucano–Tariana use classifiers with demonstratives. Unlike Tucano, Baniwa and Tariana use classifiers with verbs in the form of topic-advancing voice, as discussed in §4.2. See table 14.2.

Other North Arawak languages of the region just have numeral classifiers (Warekena, and a few North Arawak languages in Colombia), or else no classifier systems at all (Bare).

The Makú languages of this region also have several types of classifiers: generic noun classifiers – an extremely rare type in the Amazon – and locative classifiers (see Martins 1994; Aikhenvald forthcoming). Nàdebh only has possessive classifiers of a generic type. However, these are distinct from classifier types found in other languages; and may represent individual innovations in the Makú languages of the area (see §6 of chapter 9).

¹⁰ The phoneme d: is a regular correspondent of Tariana y in some dialects of Baniwa.

¹⁰ Dâu has possessive classifiers of a different kind; see §6 of chapter 9.
are genetically closely related, it is hard to distinguish areally diffused from genetically inherited features.

Dâw, Hupda and Yuhup – Makú languages spoken in the Vaupés region in Brazil – show a few instances of unilateral diffusion from East Tucano languages. Diffusion patterns from East Tucano languages into Makú include loss of prefixing (as compared to the Makú language spoken outside the Vaupés, Nadéb); acquisition of core case marking associated with topicality of the noun; and development of verb compounding which results in the creation of aspect- and valency-changing morphemes. The contact between Makú and East Tucano peoples is much older than that between the Tucano and the Tariana, and this might explain the fact that – unlike in the case of Tariana – areal diffusion has affected the word structure. However, the degree of cultural integration between East Tucano and Makú is much less than that between Tariana and East Tucano. This may account for fewer signs of diffusion than in the case of the Tariana–Tucano contact. This unilateral character of diffusion is accounted for by the fact that the Makú used to know at least one Tucano language, but not vice versa. However, more work is needed on both Brazilian and, especially, Colombian Makú languages such as Kakua and Nukak (spoken outside the Vaupés) to distinguish between further possibilities of areal diffusion and genetically inherited developments in this family.

Tariana and East Tucano languages have been in contact for no more than about 400 years. The settlement of East Tucano tribes on the Vaupés goes further back (cf. Nimuendajú 1982: 169–70); we do not have any exact dates. The other reasonably well-described linguistic areas of the world, e.g. the Balkans, Arnhem Land in Australia (see Heath 1978, 1981), Mesoamerica (Campbell, Kaufman and Smith-Stark 1986), South Asia (Masica 1976) and the linguistic areas of North America north of Mexico (Shetzer 1976), such as the north-west coast, are considerably older than this.

As I have argued elsewhere (Aikhenvald 1996c), a study of types of Tariana placenames shows that two of these types of names are predominantly monolingual – 'historical' names which refer to places where the Tariana used to live in the remote past, and 'mythological' names which refer to the adventures of characters in origin myths. In contrast, placenames which refer to actual dwelling sites are multilingual, and are usually calqued into several languages. Even when 'historical' placenames also have names in languages other than Tariana they are never calque translations from one language into another. These properties of 'historical' and 'mythological' placenames, unexpected in an environment of obligatory multilingualism, suggest that the Tariana might have arrived in the Vaupés from a predominantly monolingual context, and that they have adopted multilingualism fairly recently.

The Vaupés linguistic area can be compared – both in time depth and in degree of multilingualism – to Kupwar village in India. According to Gunnerz and Wilson (1971: 153), the coexistence of Urdu, Marathi and Kannada in the region goes back about three or four centuries, when the Urdu-speaking Muslims arrived there. However, Kannada-speaking and Marathi-speaking people have been in the region for more than six centuries. Unlike Kupwar, the Tariana–Tucano contact area has not reached the same degree of morphosyntactic isomorphism between genetically unrelated languages, which may have been a partial consequence of a 'puristic' tendency to keep languages as much apart as possible. Language attitudes in the Vaupés, which exclude lexical borrowing and lexical diffusion, are also different from the situation in the Kupwar (where occasional lexical borrowings do take place).

The existence of structural and even formal similarities shared by Tariana, East Tucano languages and Baniwa – but absent from other North Arawak languages – shows a certain amount of diffusion in an area which goes beyond the Vaupés into the basin of Içana and its tributaries (see map 13).19

BIBLIOGRAPHY


19 Traditional stories of the Tariana and Baniwa indicate that these peoples had been in contact with the East Tucano groups before the Tariana had moved from the Rapids of Apui on the Aleri river (a tributary of the Içana) to the Vaupés. A story of the actual move from the Apui called myaka-dapanu (old CL: HABITAT) the dwelling of the ancients' describes how the Tariana lived together with the Cubeo before they left Apui.

Further studies are needed to investigate the ancient contacts between Tucano and Baniwa in the Içana–Vaupés area. For the time being, one can hypothesize that this contact must date back to a time preceding the creation of the Tariana–Tucano diffusion area in the Vaupés.
14 Içana-Vaupés linguistic area


The Upper Xingu as an incipient linguistic area

LUCY SEKI

1 THE UPPER XINGU AND ITS LANGUAGES

When the Portuguese came to Brazil they lost no time in colonizing the coasts and major rivers. The Xingu, a major southern tributary of the Amazon, is navigable for only 200 kilometres. After this there are innumerable rapids, which proved a barrier to the invader. As a consequence, the Upper Xingu area remained unmo­lested. The Indian tribes living there were able, by and large, to maintain their traditional way of life. Other tribes converged on this region, as a haven of escape from the white man's takeover of Brazil. (In addition, some tribes were sent to the Xingu area by the Europeans.) Contact with the outside world was only really established in the 1950s.

As a result, the Upper Xingu region - which was, in 1961, declared a 'national park' by the Brazilian government -- is remarkable for its linguistic diversity. Cultural traits diffuse more quickly than linguistic ones. The Upper Xingu has become culturally rather homogenous, due to close intertribal contacts, a network of intermarriage and the sharing of material culture. But linguistic diffusion is as yet in its early stages; what we have here is an incipient linguistic area.

1.1 Languages spoken

There are currently 17 indigenous groups in the Upper Xingu and the adjacent areas. See map 14. Only 10 of them² have been there for more than 100 years. The languages spoken and the number of speakers³ in the Upper Xingu are listed in

---

1 I am grateful to Frantomé B. Pacheco, for information on Carib languages.
2 Since Nuhukwa and Mutipu live together in one village, they are sometimes considered as one group (Seki forthcoming).
3 Numbers of speakers in tables 15.1-15.3 are given according to Baruzzi, Rodrigues, Mendonça et al. (1995); data on Yawalapi are from Mujica (1992).
There are 3 tribal groups speaking Arawak languages. Waurá and Mehinaku are dialects of one language. Yawalapiti shares 80 per cent vocabulary with Waurá-Mehinaku but the grammar is very different, so that there is no mutual intelligibility and Yawalapiti has to be considered a separate language.

Kuikúro, Kalapalo, Nahukwá and Matipu are mutually intelligible, and constitute a single language of the Carib family; this is markedly different from the other Carib languages spoken south of the Amazon (Franchetto 1995: 53).

Kamaiurá belongs to the Tupí-Guaraní subgroup of the Tupí family, while Awetí is a separate, single-member family within Tupí (see chapters 4 and 5 on Tupí and Tupí-Guaraní).

Materials available on Xinguan languages vary in quantity and in quality. A full grammar has been produced for Kamaiurá (Seki forthcoming). There are good materials on Trumai (Guirardello 1992, forthcoming), Kuikuro (Franchetto 1986, 1995) and Waurá (Richards 1973; 1977; 1988; 1991). Partial descriptions are available for Mehinaku (Medeiros 1990) and Yawalapiti (Mujica 1992); Awetí remains the least known of all (there is just a description of the phonology in Emmerich and Monserrat 1972). This makes difficult a full study of grammatical and lexical diffusion within the area.

Table 15.2 shows the language groups spoken adjacent to the Upper Xingu region; two of them belong to the Jê family, and one is Tupí. Their contacts with the peoples of the Upper Xingu have been comparatively recent.

### 1.2 Historical background

The Arawak-speaking peoples are believed to have been the earliest settlers in the Upper Xingu region. Other groups began to arrive from the seventeenth century – the Carib tribes, the Aweti, the Kamaiurá (Agostinho da Silva 1993: 283). According to Cowell (1973: 227), the Kamaiurá were one of the last tribes to arrive in Xingu. The Trumai are the most recent arrivals. The German explorer Karl von den Steinen established the first contact with Xingu peoples in 1884. He reported that the Trumai still maintained cultural differences from their neighbours (see note 8).

The Xingu culture area must have been established in the second half of the eighteenth, or in the early nineteenth, century (Heckenberger 1996).4

4 Extinct languages in the Xingu basin include: Tsuva and Narawâe (Carib), Kuasterâi (Arawak), and Anumanáa and Manitsawá (Tupi).

5 The archaeological evidence shows that the occupation of the basin of the Upper Xingu started in the early eleventh century. The settlement of the Lower Culture continued until at least the end of the thirteenth century (Bequellin 1993: 228). However, it is not clear what groups were involved.
Table 15.1  People and languages of the Upper Xingu

<table>
<thead>
<tr>
<th>Name</th>
<th>Family</th>
<th>Speakers</th>
<th>Traditional territory (rivers)</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a Waurá</td>
<td>WR</td>
<td>226</td>
<td>Lower Batovi; formerly Culuene</td>
<td>From the northwest; the first settlers in the region</td>
</tr>
<tr>
<td>1b Mehinaku</td>
<td>MH</td>
<td>146</td>
<td>Tuatuari</td>
<td></td>
</tr>
<tr>
<td>2 Yawalupiti</td>
<td>YW</td>
<td>13</td>
<td>Tuatuari</td>
<td></td>
</tr>
<tr>
<td>3 Kuikuro</td>
<td>QK</td>
<td>343</td>
<td>Culuene and its tributaries</td>
<td>Came from the west of the Culuene, starting from the seventeenth century</td>
</tr>
<tr>
<td>Matipu</td>
<td>MN</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kalapalo</td>
<td>KL</td>
<td>326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nahukwá</td>
<td>MN</td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Kamaurú</td>
<td>KZ</td>
<td>303</td>
<td>Lake Ypavu</td>
<td>Migrated from the north (Galvão 1953), or the north coast of Brazil (Münzel 1971: 9-10), via the Auaíá-Miqui river, through the lower Culuene where they had first contacts with the Waurá</td>
</tr>
<tr>
<td>5 Aweti</td>
<td>AE</td>
<td>93</td>
<td>Middle and upper Tuatuari and Kuliseu</td>
<td>Arrived from the west</td>
</tr>
<tr>
<td>6 Trumai</td>
<td>TU</td>
<td>100+</td>
<td>Formerly the Culuene, at present in the centre of the Upper Xingu region</td>
<td>Arrived in the Xingu region from the southeast (Araguaia and Xingu rivers), fleeing from Xavante, in the nineteenth century (Murphy and Quain 1955: 8)</td>
</tr>
</tbody>
</table>

Table 15.2  Peoples and languages adjacent to the Upper Xingu

<table>
<thead>
<tr>
<th>Name</th>
<th>Family</th>
<th>Speakers</th>
<th>Traditional territory (rivers)</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 Suyá (Mekísedje)</td>
<td>SU</td>
<td>213</td>
<td>Suyí-Missu</td>
<td>Around 150 years ago entered Xingu down the river Romaro, coming from the east</td>
</tr>
<tr>
<td>8 Txukaramae (Méuktire)</td>
<td>MT</td>
<td>577</td>
<td>The west bank of the Xingu, above the Von Martius rapids; in the 1970s moved closer to Diasarum</td>
<td>Originally a section of the Kayapo (Northern Jé), descended from the Tocantins, reaching Xingu in the mid nineteenth century</td>
</tr>
<tr>
<td>9 Juruna (Judja)</td>
<td>JN</td>
<td>181</td>
<td>The west bank of Xingu, Manitsawá</td>
<td>First recorded in 1640 close to the mouth of the Xingu; later moved south</td>
</tr>
</tbody>
</table>
The peoples recently transferred to the Xingu basin

<table>
<thead>
<tr>
<th>Name</th>
<th>Family</th>
<th>Speakers</th>
<th>Present territory</th>
<th>Traditional territory</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Rapen (Tukú)</td>
<td>TX, Carib</td>
<td>214</td>
<td>Left bank of the Xingu, close to the Uavi river, near Tromai, since the 1970s</td>
<td>Originally part of the Arara (a Carib group in the Itiru river basin) came to the lower Xingu via the Teles Pires river</td>
</tr>
<tr>
<td>11 Kayabi</td>
<td>Tupi-Guarani</td>
<td>655</td>
<td>Transferred to the north of Xingu in 1955-60</td>
<td>Descendants of southern Koppo, settled in the Paraquari area, to the south of the Xingu, the Parana River</td>
</tr>
<tr>
<td>12 Patamú (Kamayurá, PZ, Jé)</td>
<td>Krenakáre</td>
<td>159</td>
<td>Transferred to the Xingu in 1974, lived near the mouth of the Tapayuna, later moved to the left bank of the Xingu, then settled close to the left bank of the Xingu</td>
<td>Descendants of southern Koppo, settled in the Paraquari area, to the south of the Xingu, the Parana River</td>
</tr>
<tr>
<td>13 Tapayuna (East, Suyá, Branco do Piná, Novo Suyá)</td>
<td>TP, Jé</td>
<td>58</td>
<td>Transferred to the Xingu in 1969-70, settled close to the Tapayuna</td>
<td>The Aivos and Sangue rivers</td>
</tr>
</tbody>
</table>

1.3 The mobility of groups, language loss and further contacts

An important problem in the Xingu area is the frequent change of location by tribal groups. Due to intertribal hostilities and warfare, some groups become weaker; they often end up living with another group for protection. This frequently results in language endangerment and loss. In the past, the Kuyapu (Arara group) came to live with the other Arara groups, Waurá and Mehinaku; Anumaní and Tapiu (Tupi) joined Awaí (Tupi); and Manitsawá (Tupi) joined Suyá (Jé). Kamaiurá, Anumaní and Manitsawá are now extinct.

The oral history of Kamaiurá provides another example of possible language loss in the past. These people say that the present-day Kamaiurá is the result of five dialect (or language) groups coming together, for unknown reasons; currently only one person is considered to be a 'real' Kamaiurá. Other Kamaiurá, when asked, maintain that 'they are stealing the Kamaiurá language'.

Recently, the Nahuká came to live with the closely related Matipu (both Carib). In the Lower Xingu, the Tapayuna joined the Miti, both Jé, for reasons of protection. The Trumai have changed their location several times during the past 100 years, moving from the lower Culuene up to the northern parts of the upper Xingu, and then to the middle Xingu (Guirardello 1992: 12).

The mobility of groups within the Upper Xingu results in increasing contacts between them, and also in the integration of groups which did not originally belong to the Upper Xingu. The expansion of white culture brought the peoples of the Upper Xingu into closer contact with some of their neighbours. Suyá (Jé) are a case in point. They have never been part of the Upper Xingu region, but as the result of interactions with the Xinguans, they have absorbed a number of cultural traits they did not have before, e.g. hammocks, canoes, manioc (Seeger 1974, Galvao 1953). These ties have become stronger since 1959, when regular contacts with white people were established. As a result, the Suyá are considered marginally representative of the Upper Xingu culture area.

Several groups which traditionally did not inhabit the Xingu basin were transferred to the Xingu National Park by the Brazilian government to save them from extinction and to free their territories for exploration. Peoples who have been transferred into the Xingu basin since the 1950s are listed in table 15.3. In the long run, the presence of new peoples in the Xingu area and their mobility within the region must have an effect on the more traditional inhabitants of the Upper Xingu.

2 THE UPPER XINGU AS A CULTURE AREA

The peoples of the Upper Xingu share a number of distinctive cultural and material traits developed as the result of long-term contact between the groups, intertribal
marriage and economic exchange. The main material traits include haystack-shaped houses arranged in a circle around a central plaza; dependence upon fish rather than game for protein and the lack of fermented drinks. All the peoples in the Upper Xingu have benches in the form of animals fashioned from a single piece of wood, whistling arrows, bark canoes and the bull-roarer. They use the spear-thrower in intertribal games and ceremonies. Men cut hair in a circular line above the ears, while women wear it cropped on the front in a straight line. Xinguan women wear ulari, a triangular bark pubic cover. Typical Xinguan necklaces are made of rectangular or disc-shaped pieces of periwinkle shells (Galvão 1953, Murphy and Quain 1955).

There are strong similarities in social organization and kinship systems. All the Xingu groups have extended families which are predominantly patrilineal; a wife will normally move to her husband's village, after an initial period in her own village. "Common" people (camara) are differentiated from a small elite which consists of tribal leaders.

The peoples of the Upper Xingu share cultural practices. Shamanism is much the same among all groups. Sun and Moon feature as the main creative forces in their mythology. They share the main festivities, for instance, kwarry (a term of Kamaiurá origin), the feast of the dead; a martial game called huka-huka; and the dance ritual of jawari. The Xinguan peoples also have dances involving wooden or straw masks, and ceremonial flutes which are not allowed to see (Galvão 1953).

Alongside cultural uniformity, each group has its own speciality. An important part of the traditional life in the Upper Xingu is intertribal meetings (known as mottara) for the purpose of exchanging products supplied by individual groups. The Kamaiurá produce black bows, the Trumai make traditional stone axes, Carib groups are responsible for necklaces made of periwinkle shells, while the Waurá make ceramic pots. Trumai and Mehinaku provide salt (which was traditionally extracted from aquatic plants). Aweti and Mehinaku traditionally played the role of middlemen.

3LINGUISTIC SITUATION

In spite of the cultural similarity, each group maintains its identity – its language is its main distinctive feature. The idea of a language as a symbol of identity is supported by the restrictions on language use during ceremonies: one is not allowed to use a language/dialect other than one's own in the situation of intertribal communication and ceremonies. Interaction takes place via a system of non-verbal communication shared by all the participants (Emmerich 1984).5

Monolingualism is not rare among the peoples of the Upper Xingu. Many people, however, know, or at least understand, more than one language due to intertribal marriages. Since, before a couple moves to the husband's village, they tend to stay in the wife's village for a while (Galvão 1953), both have an opportunity of learning each other's language, and the children learn their mother's and their father's language. However, exogamy is not required (unlike in the Vaupés: see chapter 14). Generally speaking, the linguistic situation in the Upper Xingu can be characterized as that of "passive bilingualism" (Emmerich 1984, Basso 1973). Yawalapiti and Trumai are the only two multilingual groups.

By the end of the nineteenth century, the Yawalapiti were already a small group. They did not have a village of their own but lived among the Kuikúro, Waurá, Mehinaku, Aweti and Kamaiurá. In 1950, the members of the expedition Roncador Xingu convinced them to build their own village (Cowell 1973: 227), and their numbers increased. This was also due to intertribal marriages, and resulted in the integration of numerous speakers of Kamaiurá and Kuikúro into the Yawalapiti community (Viveiros de Castro 1977: 69; Mujica 1992). As a result, the linguistic situation in the Yawalapiti village is more complicated than elsewhere in the Xingu. The Yawalapiti language continues to mark the tribal identity; however, every Yawalapiti knows another language. Mujica (1992) reports that most children have at least one non-Yawalapiti-speaking parent, and they prefer to speak a language other than Yawalapiti. As the result of this, Yawalapiti is endangered – of the 130 inhabitants of the Yawalapiti village, only 13 are fluent in the language (1992: 7).

Similarly, the Trumai4 have been in decline since the end of the nineteenth century (see Murphy and Quain 1955: 9f.). In 1938, many were bilingual in Kamaiurá. Nowadays, all the Trumai speakers are multilingual, and they tend to know more

4 Basso (1973: 5) described the Upper Xingu as a 'communication network' rather than a 'speech community', i.e. "a system consisting of several kinds of linkages between individu-
Portuguese than do other Xinguan peoples. Children tend first to learn languages other than Trumai (claiming that Trumai is too hard). Trumai is likely to become endangered in the near future (Guirardello 1992: 20).

4 INCIPIENT AREA DIFFUSION IN THE UPPER XINGU

Given the short time depth and limited multilingualism, one would not expect to find many traces of linguistic diffusion in the Upper Xingu.

Since the Arawak peoples were probably the first to have settled in the area – being followed by Carib and Tupí, and then by Trumai – one would expect the greatest areal diffusion to be between Carib, Arawak and Tupí, and from these languages into Trumai.\(^9\) The \(\text{J}e\) groups listed in table 15.2, and Juruna (Tupí), are marginal to the Upper Xingu, and one would not expect to find much areal diffusion into these languages.

There is evidence of lexical borrowing, mostly from Kamaiurá into Trumai and into Yawalapiti. According to Murphy and Quain (1955: 8–9), "the songs of two major Trumai ceremonies, both of which are practiced also by Kamaiurá, show a preponderance of Kamaiurá words and phrases", while words of Carib and Arawak origin 'are frequent in other songs'. Among lexical loans from Kamaiurá into Trumai are \(\text{tunawá} 'mat'\), \(\text{y'a} 'cabbage'\), \(\text{janý} 'pequi oil'\), and \(\text{wryaný 'vulture'}\). According to Franchetto (1986: 126) Kuikuro myths and ceremonial discourse are reported to bear an Arawak influence.

There is a certain amount of lexical diffusion from Kamaiurá into Suyá, a marginal member of the Upper Xingu culture area, e.g. \(\text{Suyá kani} 'mixture of water and manioc bread'\), \(\text{parana} 'river'\), \(\text{wonui} 'corn'\) from Kamaiurá \(\text{kaví}\), \(\text{parana}, \text{awasi}\) (cf. Seeger 1978: 161).

Some languages of the Upper Xingu share certain traits which are not found in genetically related languages outside the area and may have developed as the result of areal diffusion. Four of these traits are:

(A) The development of the vowel phoneme \(i\) in Yawalapiti

Vowel systems in all the Xinguan languages include \(i, a, i\) and \(u\).\(^{10}\)

The following chart shows vowel correspondences between proto-Arawak, Waurá and Yawalapiti (Seki and Aikhenvald forthcoming). The phonological changes in

---

\(^9\) The Trumai, being the latest arrivals (and among the weakest tribes in the Upper Xingu), would not be expected to be the source of diffusion and borrowing. Interestingly, a Trumai word \(\text{mutu} 'water'\) is used in a number of names of rivers in the northern part of the Upper Xingu, such as \(\text{Autaí Miču}, \text{Suya Miču}, \text{Maníssu-Miçu}\).

\(^{10}\) The vowel \(e\) is found in Waurá–Mehinaku, and Kuikuro. Aveli and Kamaiurá have both \(e\) and \(o\).

---

<table>
<thead>
<tr>
<th>proto-Arawak</th>
<th>Waurá</th>
<th>Yawalapiti</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>i</em></td>
<td><em>i</em></td>
<td><em>i</em></td>
</tr>
<tr>
<td><em>e</em></td>
<td><em>e</em></td>
<td><em>i</em></td>
</tr>
<tr>
<td><em>i</em></td>
<td><em>i</em></td>
<td><em>i</em></td>
</tr>
</tbody>
</table>

pre-Yawalapiti have to be ordered as follows: \(i \rightarrow i; *e \rightarrow *i\). In Yawalapiti the two high vowels merged; then \(e\) shifted to \(i\). (See table 15.4.) The recreation of the vowel \(i\) (after the original \(i\) developed into \(i\)) was possibly due to areal pressure, since \(i\) is present in every other Xinguan language.

(B) Shift \(p \rightarrow h\): from Arawak to Carib and Tupí-Guaraní

A feature shared by genetically unrelated languages in the Upper Xingu region is the shift \(p \rightarrow h\). In Kuikuro, \(p\) has become \(h\) in intervocalic position (Franchetto 1995: 55). Kamaiurá (chapter 5) and proto-Tupí-Guaraní \(*pw\) has become \(hw\) or \(h\). In Yawalapiti, \(pi- '2sg prefix' \) becomes \(hi-\) if the following root begins with \(w\) or \(y\). This change is not found in other Tupí-Guaraní or Carib languages. It is not frequent in Arawak languages outside the region. For example, \(p\) has become \(h\) in Pareí, the Arawak language which is genetically and geographically closest to Xinguan Arawak, and also in a number of North Arawak languages (e.g. Bahiana, Achagua, Yaúteró). One may hypothesize that the source of this diffusional feature could have been the Arawak languages.

(C) CV syllable structure: from Arawak into Carib

All Arawak languages, including those spoken in the Upper Xingu, have CV syllable structure. Carib languages spoken outside the Upper Xingu have CV and CVC syllables, while the Xinguan Carib language (with its four dialectal varieties) has only CV syllables. It may well be that this feature is the result of Arawak influence.

(D) Loss of genders in Xinguan Arawak: a feature diffusing from Carib and Tupí into Arawak

The Arawak languages of the Upper Xingu have lost the gender distinction between masculine and feminine in cross-referencing markers. Waurá-Mehinaku has also lost genders in independent pronouns. We may hypothesize that this loss could have occurred as the result of contact with Tupí and Carib languages in the Xingu, none of which have genders.

These traits are indicative of an early stage of areal diffusion in the Upper Xingu region. The exact amount and the character of lexical borrowings and other
diffusional features in the languages of the Upper Xingu require further studies, both descriptive and comparative.

5 CONCLUSIONS

The Upper Xingu culture area is relatively young - 150-200 years old. The time depth has not been long enough for a linguistic area to develop; this is the reason why, in spite of cultural similarities and a certain degree of multilingualism, 'strong linguistic indicators of long-term mutual influence are not abundant here' (Campbell 1997: 340). However, the existing multilingualism, lexical borrowings and a few shared structural features are suggestive of an incipient linguistic area.

This illustrates well the problems which might be caused by assuming 'that culture areas and linguistic areas will coincide' (Campbell 1997: 340). The Upper Xingu area is thus reminiscent of the much-discussed case of the Great Plains region of the United States - a recognized culture area, but not a linguistic one. Similarly to the Great Plains, the Upper Xingu area has developed recently; in both the Upper Xingu and the Great Plains there has not been sufficient time depth (and, possibly, social stability) for a true linguistic area to be established (Doug Parks p.c.; see also Sherzer 1973: 773; Bright and Sherzer 1976: 235).

Contacts with white people are now increasing in the Upper Xingu region. In the 1970s a major highway was constructed through the National Park. Indigenous neighbours who did not originally belong to the Upper Xingu culture complex are undergoing strong influence from Upper Xingu culture. This goes together with the mobility of tribal groups who freely move about the whole Xingu National Park. The scope of what used to be the Upper Xingu culture area is thus increasing, and the situation is less and less stable. Thus, it is hard to predict whether recent social developments will lead to further areal diffusion, and what direction this diffusion might take.

BIBLIOGRAPHY


Seki, L. and Aikhenvald, A. forthcoming. 'Arawak languages of Xingu: a reconstruction'.


INDEX OF LANGUAGES AND LANGUAGE FAMILIES

<table>
<thead>
<tr>
<th>Language</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>A'ukré, see Kayapó</td>
<td>207</td>
</tr>
<tr>
<td>A?uy?k?e'te'ku'talk?e'su'</td>
<td>269</td>
</tr>
<tr>
<td>A?ulan'te'su'</td>
<td>269</td>
</tr>
<tr>
<td>Achkogba</td>
<td>70, 72, 427</td>
</tr>
<tr>
<td>Ačê, see Guayakí</td>
<td>307, 312, 321, 323, 325, 327-8, 331</td>
</tr>
<tr>
<td>Achkua, see Achkua</td>
<td>307, 312, 321, 323, 325, 327-8, 331</td>
</tr>
<tr>
<td>Achuar</td>
<td>307, 312, 321, 323, 325, 327-8, 331</td>
</tr>
<tr>
<td>Achuar-Shiwiar</td>
<td>309</td>
</tr>
<tr>
<td>Aqogun, see Aguaruna</td>
<td>307-9, 312-14, 332, 334, 370</td>
</tr>
<tr>
<td>Aîkanà</td>
<td>341, 343, 362-3</td>
</tr>
<tr>
<td>Ajarí, see Wayoró</td>
<td>129, 131, 132</td>
</tr>
<tr>
<td>Akawaí</td>
<td>24-5</td>
</tr>
<tr>
<td>Akroí</td>
<td>167</td>
</tr>
<tr>
<td>Akuriyó</td>
<td>26</td>
</tr>
<tr>
<td>Akwáwa</td>
<td>131</td>
</tr>
<tr>
<td>Algonquin</td>
<td>11</td>
</tr>
<tr>
<td>Almosan-Keresiouan</td>
<td>327</td>
</tr>
<tr>
<td>Amalicore, see Ñiquiuto</td>
<td>108-10</td>
</tr>
<tr>
<td>Amanaye</td>
<td>132</td>
</tr>
<tr>
<td>Ampari</td>
<td>132, 141, 143-4, 149, 157</td>
</tr>
<tr>
<td>Amarákeeri</td>
<td>307, 311-12</td>
</tr>
<tr>
<td>Amuririana</td>
<td>71, 387</td>
</tr>
<tr>
<td>Amawaca</td>
<td>229-30, 233, 240</td>
</tr>
<tr>
<td>Amondawá</td>
<td>132</td>
</tr>
<tr>
<td>Amuesha</td>
<td>68, 73-4, 90-1</td>
</tr>
<tr>
<td>Amurí</td>
<td>132, 139, 145</td>
</tr>
<tr>
<td>Anauya</td>
<td>70</td>
</tr>
<tr>
<td>Andoa</td>
<td>309</td>
</tr>
<tr>
<td>Andoké</td>
<td>344, 370-2, 374-7</td>
</tr>
<tr>
<td>Añuñ, see Parauanho</td>
<td>171, 172</td>
</tr>
<tr>
<td>Apaful</td>
<td>24-5, 43, 140</td>
</tr>
<tr>
<td>Apó</td>
<td>127, 132</td>
</tr>
<tr>
<td>Apinaye</td>
<td>167, 171-2</td>
</tr>
<tr>
<td>Apolista</td>
<td>67</td>
</tr>
<tr>
<td>Apuriná</td>
<td>13, 68</td>
</tr>
<tr>
<td>Arabela</td>
<td>307, 309, 312, 317-18, 320-2, 328-31, 333</td>
</tr>
<tr>
<td>Arakú</td>
<td>71</td>
</tr>
</tbody>
</table>

Caribbean Arawak, see Caribe, see Garifuna |
Carijona | 24-5 |
Carirí | 71 |
Carmelo, see Yaté | 129-30 |
Caruru, see Baniwa of Icarai | 307-9, 312, 321, 323, 325, 327-8 |
| Cataó, see Baniwa of Icarai | 229-30 |
| Cashibó | 229-30, 243 |
| Cashinawa | 229-30, 243 |
| Catapothi, see Baniwa of Icarai | 229-30, 243 |
| Caujé | 344, 365-9 |
| Cavanuez | 309 |
| Castiló | 67 |
| Cayuvaí | 344, 367, 369 |
| Central Tucano | 207, 211, 219, 386, 389, 406 |
| Chaco | 229-32, 240-1 |
| Chamá, see Ese Eja | 437 |
| Chamicuro | 68 |
| Chandinawa, see Sharana | 229-30, 243 |
| Chané (Arawak) | 67, 79, 81, 88 |
| Chané (Tupí-Guarani) | 129-30 |
| Chapacura | 10, 14, 343, 343, 358-61 |
| Chapara | 310 |
| Chayahuita | 207-9, 313, 320-7, 330, 332-34, 345 |
| Chebero, see Jebero | 129-30, 370 |
| Chiqui | 129, 370 |
| Chiquin, see Tâkói | 227 |
| Chimane | 227 |
| Chiquitano | 364 |
| Chiquito | 106 |
| Chiriana, see Bahwa | 127, 130, 133, 159 |
| Chiriguano | 229-30 |
| Chiripí, see Nandéva | 129-30 |
| Chiripuño, see Arawac | 229-30 |
| Chitonawá | 307, 309, 312, 320, 325 |
| Chiru, see Jivaro | 370, 394, 406 |
| Cholón | 311 |
| Chontalpí | 68 |
| Cinta-larga | 109 |
| Cohán-Mákú | 251 |
| Comau, see Cashibo | 171-2 |
| Combó, see Shipibo | 171-2 |
| Coroado | 168 |
| Coropó, see Koropó | 207, 212, 218-19, 222, 386-91, 406 |
| Culeá | 261, 344, 370-7 |
| Cullina, see Kuna | 229 |
| Cuní | 68 |
| Cushiní | 68 |
| Dáw | 252-5, 257, 259, 386, 394, 398-400, 404, 406-7, 410, 412 |
Mabanaro 344
Macuá 344
Machiguenga 13, 68
Machoto, see Ionuma
Machito-Piro 26, 167-170, 173, 176, 179, 182, 186, 188-90, 192, 194, 197
Mba'anongo 217, 127, 130, 140, 171, 173, 176, 179, 182, 186, 188-90, 192, 194, 197, see also Pipó
Mayan 11-12
Mayna, see Ionuma
Máyora 68, 229
Mbabaram 11
Mbbýa 127-8, 130, 138, 140-1, 143, 153, 158, 189
Mehinaku 67, 419-20, 423-7
Meköns 10-9
Mëkëndjë, see Suyá
Mënö 167
Menkangnom, see Kayapó
Menktité, see Kayapó
Merrine 165
Mëtëkëntë, see 421, 423
Mincke Witoto 310
Minëyàyräm, see Krenák
Minëra 310, 312
Mundë, see Aikana
Munduruku 95, 107-10, 112-20
Munich 307, 311
Munsibori, see Munichi
Münkä, see Irantxe
Murta 353-5
Murà-Pirãhd 4, 341, 343, 353-7
Muraro, see Candoshi-Chapera
Murut Wioto 307, 310, 316-17, 321-2, 326, 331
Mynky, see Irantxe
Nabëd 9, 45, 252-5, 257, 260, 376, 387, 394, 396, 398-400, 404-6, 410, 412
Nahu, see Yoranawa
Nahuajka 26, 419, 20, 423
Nakarróthe 269-70
Nakanyánuk, see Krenák
Nakpé, see Krenák
Nakpibé, see Krenák
Nakyañib, see Kunak
Nambubara 12, 269-91
Nambuyara, see Nambyua
Naná 307, 311
Namyala, see Yoranawa
Nayaramú, see Mundë
Nèçu 269
Nèndëhâ 127-8, 130
Nëyândji, see Krenák
Nipês 310, 317
Nokamàn 229
Nontocoëngas 13, 68, 79
North Arakan 9, 69, 74-81, 84-8, 90-5, 97-101, 258, 260, 264, 373, 385-7, 390, 392-6, 399-401, 403-7, 410-11, 413, 427
Nögëpad, see Káro
Nuká 252-5, 386, 394, 412
Nyénpëy, see Krenák
Ocaina 74, 307, 310, 316-17, 371
Olâyè 165, 168, 171, 173, 181, 188
Olâyè-Xavante, see Olâyè
Old Guarani 126-7, 130
Old Irish 11
Old Warekena 386
Omañwà, see Kokáma
Omrano, see Ofayè
Ofe, see Ofayè
Oro, see Warí
Oro, see Warí
Oro, see Warí
Oro Win 243
Otahane, see Munichi
Otë 166
Otëske 166, 168
Pacana Noawns, see Warí
Pacassa 229
Pacuí, see Baniwa of Içana
Pacë 370
Pailenanoke, see Capanawa
Paimoa 67
Pàité, see Sucir
Pajonsal Campa 68, 94
Páka Noawn, see Warí
Pakur 9, 69, 74, 91, 95, 258
Pamanahuri 309
Panará 167, 169, 171, 181, 190, 422
Pare-24
Pásara 207
Pañahame, see Kaposo
Pano family 12, 61, 95, 227-48, 364-5
Pano language 227, 229
Parraní, see Wirapano
Papià 207
Papuan 405
Paraguayan Guarani 127-8, 130
Parakanã 131
Parâna, see Kaingang
Paraulhão 69, 89
Parœie 67, 81, 427
Pareci-Caxiúba 67
Pareci-Saraveca 67
Pareci-Xingu 67, 80, 83, 86, 90, 93
Parianto 132, 136, 138-9, 141, 148, 151
Parqunahuwa, see Yoranawa
Passe 71
Passe 387
Patahö, see Pataxó
Pataxó 165, 167
Paujile, see Baniwa of Içana
Pawamani 12, 292-306
Pauna 67
Paweti 132
Pawishiana 26
Pawumwan 359
Peba 310, 312
Peba-Gyua 307, 310, 319, 321, 327
Pedra Branca 108
Pemong 24-5
Pereñe 68
Peruvian Arawak 10, 93
Phipoço 70
Pirão 344, 370
Pichi 68
Pinche, see Taushiro
Pirahã 354-5
Piriatupys 207, 209, 214, 386-8, 396-7
Piro, see 13, 68
Piro-Apuirína 68, 80, 90-1
Pisimãra 209
Piyana 344
Piyãn, see Krenák
Portuguese xxv-xxx-xxv, 1-2, 7, 19-20, 23, 72, 125, 127, 140, 157, 177, 184, 251, 254, 271, 292, 310, 345, 358, 388, 391, 426
Potim, see Krenák
Poturu, see Zoé
Poyawna 229
Pre-andine Arawak 16, 10, 312
Pre-Jamamadi 14
Pro-Amerindian 14
Proto-Arawá 12-14, 295-6, 298, 300, 304
Proto-Arawã-Arawak 12
Proto-Arawak 12-14, 74-5, 77-81, 84-5, 88, 93, 393, 403, 426-7
Proto-Asháninka 13
Proto-Caribo 36, 31-2, 37, 45, 60
Proto-East-Tucano 393
Proto-Jê 165, 172, 178
Proto-Lokono-Guarajó 74
SUBJECT INDEX

absolutive 61, 96, 155-7, 187, 194, 241-2, 320, 348, 366-7
accusative 9-10, 32, 35-7, 60-1, 90, 118, 121, 187, 241-2, 254, 263
active verbs xxvi, 32, 81, 83, 85-7, 117, 146-7, 154-5, 157-60, 213, 351, 407
adverbialization 57-8
ambitransitive verbs xxvi, 81, 86, 91, 259, 300, 349, 355
applicative xxvii, 9, 80, 91-2, 262, 301, 321, 327, 334, 375
aspiration 77, 79, 201, 295, 313
auxiliary verbs 19, 55, 60, 89, 97-8, 213, 216, 235, 247, 301, 366, 379, 405
body parts 9, 41, 45, 82, 85, 95, 116, 149, 152, 218, 243, 298, 330, 346-9, 353, 363, 369, 399
causative xxvii, 9, 44, 80, 86, 90-2, 98, 120, 158, 195-6, 244, 247, 260-1, 301, 316, 327-8, 347, 349, 353, 356, 358, 361, 369, 375, 379
classifiers xxvii, 1, 8-10, 42, 80, 82-4, 86-7, 116, 185, 218-19, 221, 225, 258, 280-1, 290-1, 319, 324, 332, 347, 350, 362-4, 366, 369, 373-4, 379, 399-403, 407-10
collective number 38-42, 53, 84, 183, 259, 349, 360, 374
complementizers 121
competitive aspect 38, 93, 214-15, 260, 301-2, 326, 379
conditional clauses 52, 278
conditional mood 10, 80, 83, 94, 96-7, 100, 116, 148, 177, 185, 192-3, 214, 221, 258, 298-9, 302, 319, 360, 362-3, 366, 369, 373, 379, 400-1, 408, 427
conceptualization 28, 76, 79, 313
dependent-marking xxvii, 10, 80, 180, 254, 346, 352, 363, 372, 396
desiderative mood 58, 93-4, 216, 245, 326
detransitivizers 31, 44, 159
different subject marking xxviii, 101, 197, 224, 405
ditransitive verbs xxvii, 154, 158, 243, 286-7, 327, 376, 379
dubitative mood 93-4, 274, 318
durative aspect 93, 301, 326
ejectives 313, 352
epenthesis 31, 136
split-ergativity 9, 32, 80, 87, 89-90, 98, 263, 348, 366, 375, 403
evidentials 3, 93, 213-14, 216, 222
evidentiality xxvii, 1, 8, 87, 119, 214, 261, 274, 277-9, 282, 286, 288, 302, 349-50, 376, 404
feminine, see gender
gender 8, 87, 10, 80, 83-4, 94-6-7, 100, 116, 148, 177, 185, 192-3, 214, 221, 258, 298-9, 302, 319, 360, 362-3, 366, 369, 373, 379, 400-1, 408, 427
glottophony 28, 76, 79, 313
habitual aspect 93, 260, 325, 349, 360
head-marking xxvii, 8, 10, 80, 114, 180, 254, 273, 346, 357, 359, 362-3, 366, 369, 372, 378, 396, 398
hierarchy, person/agentivity 33, 36, 155
imperfective aspect 274-6, 356
implosives 210, 317, 365
incomplete aspect 38, 245, 326
incorporation 9-10, 45, 92, 159, 254, 261-3, 265, 324, 363, 369, 376
Subject index

noun incorporation 9, 45, 95, 220, 243, 261, 350, 376
inverse 31-2, 34
iterative aspect 214-15, 301, 349, 356, 360
kinship terms 41, 82, 183, 218, 258, 330, 395, 399-400, 403
labialization 29, 140-1, 318
laryngealization 271
masculine, see gender
nasalization 8-9, 78-9, 134-5, 143-7, 153, 179-80, 183, 194, 246, 260, 277-80, 287, 290-1, 334, 358
negation 48-9, 52, 54-6, 93, 154, 222, 244, 264-5, 301, 349, 356, 360, 363, 367
negative imperatives 96, 276
prohibitives 96, 260, 264-5, 379
nominalizations 10, 49, 51, 56-1, 121, 144, 159-60, 183, 194, 246, 260, 277-80, 287, 290-1, 334, 358
nominalizers 40, 45, 48-50, 98, 143, 218-19, 223, 281, 300, 344-5, 356
nominative 90, 147, 187, 241-2, 254, 263, 320-1, 349, 399, 369, 378, 403-4
oblique-topicalized verbs 156
palatalization 29, 138-40, 318
perfective aspect 32, 39-40, 93, 274-6, 326, 349, 356, 379
pitch accent 79, 114, 212, 256, 313-14, 316-18, 394, 406-7
pivots 49, 51, 60, 90, 99, 101, 121-2, 195, 263, 298, 301-2, 304-6, 353
possession 48-9, 52, 58, 96, 134, 154, 222, 244, 264-5, 301, 349, 353, 356, 358, 360, 404
negative imperatives 96, 276
prohibitives 96, 260, 264-5, 379
nominalizations 10, 49, 51, 56-1, 121, 144, 159-60, 183, 194, 246, 260, 277-80, 287, 290-1, 334, 358
nominalizers 40, 45, 48-50, 98, 143, 218-19, 223, 281, 300, 344-5, 356
nominative 90, 147, 187, 241-2, 254, 263, 320-1, 349, 399, 369, 378, 403-4

reciprocal 44, 85, 88-9, 92-3, 119-20, 154, 159, 193, 237, 243-4, 260, 282, 328, 356, 360, 376
reduplication 80-1, 86, 119, 183-4, 260, 302, 304, 323, 325, 360-1, 303, 367
reflexive 37, 44, 88, 91-3, 119-20, 154, 159, 182, 195, 237, 243-4, 260, 283, 303, 328, 349, 356, 358, 360, 369, 376
relativization 265, 287-8, 332-3
relative clauses xxvi, xxviii, 59, 95, 99-100, 152, 222, 236, 245, 265, 287, 332-3, 350-1, 356, 360-1, 377
S xxvi, 32-5, 82, 86-90, 97-8, 101-2, 117, 147, 155, 158-60, 375, 403, 407
same subject marking xxviii, 88, 101, 197, 224, 405
serial verbs xxvii, 97-8, 147, 157, 213, 360-1, 363, 404
S xxvi, 32-5, 86-90, 98, 100-1, 117, 146, 155, 158, 160, 375, 403, 407
stative verbs xxvi, 32, 81, 82, 86, 89, 91, 95, 117, 146-7, 149, 151, 154-5, 160, 213, 221, 257, 284, 300-1, 346, 351, 375, 407
subordination 265, 332-3
subordinate clauses 287
tone-changing derivations 9, 90-3, 260-1, 356, 366
decreasing, see reducing below
increasing xxvii, 87, 90-1, 101, 163, 372, 375
reducing xxvii, 90, 98, 260, 367, 376
tone-changing processes, see tone-changing derivations
valency-changing derivations
valency-changing devices, see valency-changing derivations
valency-changing processes, see valency-changing derivations
verb compounding 98, 216, 259-60, 350, 356, 361, 376, 404, 411-12
vowel assimilation 256, 298, 345, 354
vowel harmony 79, 81, 232, 359
vowel length 28-9, 313-14, 362, 395
vowel shift 112, 144-5
THE AMAZONIAN LANGUAGES

Paperback Re-issue

The Amazon Basin is arguably both the least-known and the most complex linguistic region in the world today. It is the home of some 300 languages belonging to around 20 language families, plus more than a dozen genetic isolates, and many of these languages (often incompletely documented and mostly endangered) show properties that constitute exceptions to received ideas about linguistic universals. This book provides an overview in a single volume of this rich and exciting linguistic area. The editors and contributors have sought to make their descriptions as clear and accessible as possible, in order to provide a basis for further research on the structural characteristics of Amazonian languages and their genetic and areal relationships, as well as a point of entry to important cross-linguistic data for the wider constituency of theoretical linguists.