

Tense, Aspect, Modality, and
Evidentiality Marking in South American
Indigenous Languages

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: Lama at Machu Picchu by Neele Mueller

ISBN: 978-94-6093-107-9
NUR 616

Copyright © 2013: Neele Mueller. All rights reserved.
ERC Advanced Grant to Pieter Muysken, “Traces of Contact”

Tense, Aspect, Modality, and Evidentiality Marking in South American Indigenous Languages

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,

volgens besluit van het college van decanen

in het openbaar te verdedigen op dinsdag 26 maart

om 10.30 uur precies

door

Neele Janna Müller

geboren op 28 mei 1982

te Leer

Promotor: Prof. dr. Pieter Muysken

Copromotoren: Dr. Rik van Gijn (Universitat Zurich)
Dr. Mily Crevels

Manuscriptcommissie: Prof. dr. Helen de Hoop
Prof. dr. Kees Hengeveld (Universiteit
van Amsterdam)
Prof. dr. Pier Marco Bertinetto (Scuola
Normale Superiore, Pisa)

För Mama un Papa.

Für Axel.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	vi
ABBREVIATIONS AND SYMBOLS	viii
1. INTRODUCTION	1
2. TERMINOLOGY, FRAMEWORK, AND TOOLS	3
2.1 INTRODUCTION	3
2.2 TERMINOLOGY	3
2.2.1 Marking	3
2.2.2 Cumulative morphemes	5
2.2.3 Ambiguity	6
2.3 CONCEPTUAL FRAMEWORK	7
2.3.1 Prototype theory and the canonical approach	7
2.3.2 The dominance parameter	8
2.3.3 Grammaticalization	9
2.3.4 Obligatoriness	10
2.3.5 Systematicity	11
2.4 THE MODUS OPERANDI OF THIS STUDY	11
2.5 TOOLS	15
2.5.1 The language sample	15
2.5.2 The questionnaire	22
3. TENSE	25
3.1 INTRODUCTION	25
3.2 DEFINITIONS	26
3.2.1 PRESENT, PAST, and FUTURE Tense	26
3.2.2 Remoteness degrees	32
3.3 TENSELESS LANGUAGES	33
3.4 TENSE SYSTEMS	36
3.4.1 Overview	36
3.4.2 Binary split FUTURE/ NONFUTURE	38
3.4.3 Simple split FUTURE/ NONFUTURE	40
3.4.4 Binary split PAST/ NONPAST	41

3.4.5	Three-way split Tense systems	43
3.4.6	FUTURE and PAST marking	44
3.5	MAXIMUM TENSE	45
3.6	REMOTENESS DEGREES	46
3.6.1	Overview	46
3.6.2	One remoteness degree in the PAST	47
3.6.3	Two remoteness degrees in the PAST	48
3.6.4	Three remoteness degrees in the PAST	49
3.6.5	Four remoteness degrees in the PAST	51
3.6.6	Six remoteness degrees in the PAST	52
3.6.7	Remoteness degrees in the FUTURE	53
3.6.8	Remoteness degrees in the PAST and FUTURE	54
3.6.9	Discussion	57
3.7	TYOLOGICAL DISTRIBUTION IN THE SAMPLE	63
3.8	GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE	69
3.9	GENEALOGICAL DISTRIBUTION IN THE SAMPLE	73
3.9.1	Introduction	73
3.9.2	Macro-Gêan	73
3.9.3	Arawakan	74
3.9.4	Tupían	75
3.9.5	Cariban	78
3.9.6	Quechuan	80
3.10	STABILITY OF TENSE	83
3.11	SUMMARY	85
4.	ASPECT	87
4.1	INTRODUCTION	87
4.2	PERFECTIVE/ IMPERFECTIVE	89
4.2.1	Definitions	89
4.2.2	Distribution in the sample	90
4.3	HABITUAL	93
4.3.1	Definition	93
4.3.2	Distribution in the sample	94
4.4	CONTINUATIVE	95
4.4.1	Definition	95

4.4.2	Distribution in the sample	96
4.5	ITERATIVE	97
4.5.1	Definition	97
4.5.2	Distribution in the sample	100
4.6	COMPLETIVE/ INCOMPLETIVE	105
4.6.1	Definitions	105
4.6.2	Distribution in the sample	107
4.7	ANTERIOR	111
4.8	TYPOLOGICAL DISTRIBUTION IN THE SAMPLE	114
4.9	GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE	118
4.10	GENEALOGICAL DISTRIBUTION IN THE SAMPLE	123
4.10.1	Introduction	123
4.10.2	Macro-Gêan	123
4.10.3	Arawakan	123
4.10.4	Tupían	124
4.10.5	Cariban	125
4.10.6	Quechuan	126
4.11	STABILITY OF ASPECT	128
4.12	SUMMARY	129
5.	MODALITY	131
5.1	INTRODUCTION	131
5.2	REALIS/ IRREALIS	133
5.2.1	Definitions	133
5.2.2	Distribution in the sample	135
5.3	CERTAINTY/ DUBITATIVE (EPISTEMIC MODALITY)	141
5.3.1	Definitions	141
5.3.2	Distribution in the sample	143
5.4	INTENTIONAL	146
5.4.1	Definition	146
5.4.2	Distribution in the sample	146
5.5	POTENTIAL	149
5.5.1	Definition	149
5.5.2	Distribution in the sample	150
5.6	PURPOSIVE	153

5.6.1	Definition	153
5.6.2	Distribution in the sample	154
5.7	FRUSTRATIVE	158
5.7.1	Definition	158
5.7.2	Distribution in the sample	159
5.8	DESIDERATIVE	163
5.8.1	Definitions	163
5.8.2	Distribution in the sample	163
5.9	COMMAND TYPES	167
5.9.1	Definitions	167
5.9.2	Distribution in the sample	170
5.10	TYPOLGICAL DISTRIBUTION IN THE SAMPLE	180
5.11	GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE	188
5.12	GENEALOGICAL DISTRIBUTION IN THE SAMPLE	198
5.12.1	Introduction	198
5.12.2	Arawakan	198
5.12.3	Tupían	201
5.12.4	Cariban	202
5.12.5	Quechuan	204
5.13	STABILITY OF MODALITY	205
5.12	SUMMARY	205
6.	EVIDENTIALITY	207
6.1	INTRODUCTION	207
6.2	DEFINITIONS	209
6.2.1	FIRSTHAND and VISUAL Evidentiality	209
6.2.2	NON-FIRSTHAND Evidentiality	210
6.2.3	SECONDHAND and QUOTATIVE Evidentiality	210
6.2.4	INFERRED and ASSUMED Evidentiality	212
6.3	TYPOLGICAL DISTRIBUTION IN THE SAMPLE	213
6.3.1	General overview	213
6.3.2	Systems of Evidentiality marking	217
6.4	GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE	227
6.4.1	Results	227
6.4.2	The Guaporé-Mamoré area	234

6.5	GENEALOGICAL DISTRIBUTION IN THE SAMPLE	236
	6.5.1 Introduction	236
	6.5.2 Arawakan	237
	6.5.3 Tupían	239
	6.5.4 Cariban	240
	6.5.5 Quechuan	241
6.6	STABILITY OF EVIDENTIALITY	243
6.7	SUMMARY	243
7.	DISCUSSION AND CONCLUSIONS	245
7.1	INTRODUCTION	245
7.2	GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE	246
	7.2.1 Overview	246
	7.2.2 Andes vs. Amazonia	246
	7.2.3 The Chaco	248
7.3	GENEALOGICAL DISTRIBUTION IN THE SAMPLE	250
7.4	TAME PROMINENCE	256
7.5	STABILITY	258
7.6	FUTURE RESEARCH	260
	REFERENCES	263
	SUMMARY	273
	SAMENVATTING IN HET NEDERLANDS	277
	CURRICULUM VITAE	281

ACKNOWLEDGEMENTS

Acknowledging all the people who so graciously helped me during the project seems no less complex than the verb morphology in my sample, and in case I forgot to mention someone overtly, be assured that this is rather because I tend to overlook things closest to me.

I consider my thesis the verbal stem, root, or theme of this project, with the following people having scope over it: closest to the root are my supervisors, Pieter Muysken, Mily Crevels, and Rik van Gijn. Without them, the thesis would have been ungrammatical. They marked my arguments, switched references if necessary, and helped me through Ärger-tiv times. I thank Pieter especially for guiding me through the rijstepap to the rozijnen! My sincerest thanks for their belief in my abilities, for their outstanding support and patience, and for their guiding me towards becoming not only a better researcher, but also organizer, writer, typologist, etc. I could not have wished for better supervisors.

To Pieter my thanks for funding my position and my stay in Peru and giving me the possibility to embark on this four year-long journey. You let me steer the boat through the pap my own way but nevertheless guided me. How you did that I still fail to understand. Your positive look at things saved me many a time!

To Mily my thanks for making me always think twice, for minding my manners and appearance, and for the references!

To Rik, a role model on accuracy in everything he does, and for the first book present in my career. Yurakaré will always have a special place in my heart because of that!

I wish to express my gratitude to my reading committee all of whom have been part of my life as a Ph.D. student, and I am glad that they accompanied me at the last stages as well: Prof. Dr. Helen de Hoop, Prof. Dr. Pier Marco Bertinetto, and Prof. Dr. Kees Hengeveld.

Imperative for the successful verb phrase are the other project members, closest of all Olga Krasnoukhova, Josh Birchall, and Bob Borges. I think together we agglutinated to a very harmonious project complex. You have been major suffixes to me! Olga and Josh, you two are my paranymphs as well and I am very grateful for your support and friendship not only at the defense, but throughout all the years.

I thank my office mates, Mily in the first year and then Olga, Bob, and Linda van Meel for making coming to the office every day a pleasure.

I also want to thank each of the members of the group for inspirational collaboration and the occasional chat.

Sonja Gipper and Martine Bruil: it was so good to meet fellow evidentiality nerdies and I am still glad we managed to organize the “Mood and Modality” workshop!

The linguistics department in Nijmegen was always there with cake when I needed one (or not).

Ylva, Anouk, Anne, and Sten for a great time with the PON!

I am hugely indebted to many helping brains. I am immensely grateful to Loretta O'Connor for all services provided. GIS sounds a lot easier when Loretta explains it to you! The same holds for Harald Hammarström who actually managed to install statistics in my world and who helped with all kinds of computational problems.

My family and friends outside linguistics always supported me and shared my excitement.

My gratitude goes out to every single researcher whose passion served as an example to me and my thanks also go to all those who took the time to answer my queries with great detail. Their interest in languages is surpassed only by their joy of sharing it. Feel free to attach any of the descriptions (inspiring, encouraging, enlightening, instructive, invaluable) to any of the people that follow (in no special order), who where of major importance for this study: Sérgio Meira, Ana Vilacy Galucio, Margot van den Berg, Hein van der Voort, Helder Perri, Simon van de Kerke, Gerrit-Jan Kootstra, Simeon Floyd, Love Eriksen, Sebastian Drude, Pablo Irizarri van Suchtelen, Swintha Danielsen, Antje Muntendam, Kofi Yakpo, Cefas van Rossem, Hülya Sahin, Robbert van Sluijs, Ana Paula Brandão, Luca Ciucci, Christina Bergmann, Francesca Romana, Jean-Christophe Verstraete, Olesya Khanina, Fiona Jordan, Alexandra Aikhenvald, Willem Adelaar, Rodolfo Cerrón Palomino, Jorge Pérez Silva, Joan Bybee, Jeanette Sakel, Simon Overall, Sander Lestrade, Asia Kulig-Borges, Frank Seifart, Antoine Guillaume, Paul Frank, the Leipzig Summer School bunch, the Leiden Linguistics bunch, the MPI Nijmegen bunch, the Marburg bunch, and countless more who know that I mean them if they ever managed to make me smile.

Not to be forgotten because of major importance in daily office life and Ph.D matters were Hella, Dirkje, Christel, and Lisenka.

ABBREVIATIONS AND SYMBOLS

Glossings are simplified according to the Leipzig glossing rules. It was attempted to keep as close as possible to the glossing in the original while being consistent. In cases where no glossings are given in the original they were left out.

- affix boundary
- = clitic boundary
- ∅ absence of overt marker
- ? Unknown

A	A-participant; A-marking	ASSR	assertive
ABL	ablative	ATTR	attributive
ABS	absolute	AUG	augmentative
ABSN	absential	AUX	auxiliary
ACC	accusative	BEN	benefactive
ACT	active	BM	boundary marker
ADD	additive	BR	bound root
ADJ	adjective	C	coreferential
ADV	adverb(ial)	CAUS	causative
ADVZ	adverbializer	CD	classifier/ directional
AFF	affected	CERT	certainty
AFM	affirmative	CLF	classifier
AG	agent	CMT	comitative causative
ALL	allative	CN	connective
AN	animate	CNT	continuous, continuative
ANA	anaphoric pronoun	CNTRFC	counterfactual
ANC	ancient	CO	cooperative object
ANP	antipassive	COL	collective
ANT	anterior	COM	command
APPL	applicative	COMP	complement clause
ART	article	COMPL	completive
AS	absolute state	COND	conditional
ASM	assumed evidentiality	CONJ	conjunction
ASO	associative	CONT	contrast
ASPC	aspect	COP	copula

CP	comparative	FLR	filler form
DAT	dative	FM	formative
DECL	declarative	FNS	final nominal suffix
DEIC	deictic root	FOC	focus
DEL	delimiter	FRUST	frustrative
DEM	demonstrative	FUT	future tense
DEP	dependent clause	GEN	genitive
DEPM	dependent marker, topic marker	GER	gerund
DES	desiderative	GINF	generic infinitive
DEV	direct evidential	GKN	general knowledge
DEVLD	devalued	GNR	general
DIM	diminutive	GNRF	general referent
DIR	directional	GO	go, intentional, directional
DISC	discontinuative	HAB	habitual
DIST	distant, distal	HEA	hearsay
DISTR	distributive	HIAF	high affectedness Aktionsart
DR	bivalent direct	HORT	hortative
DRF	different reference	HYP	hypothetical
DSBJ	different subject	I	intermediate
DU	dual	IDR	indirect
DUB	dubitative/ doubt	IMM	immediate
DYN	dynamic	IMP	imperative
EDO	external direct object	IN	inessive relation
EMP	emphatic, emphasis	INCH	inchoative
ERG	ergative	INA	inanimate
EV	evidential	INCL	inclusive
EXCL	exclusive	INCP	incompletive
EXTRAL	extralocal	IND	indicative
EYE	eyewitness evidentiality	INFR	inferred evidentiality
F	feminine gender	INFL	inflectional
FAC	factual	INST	instrumental
FCT	factitive	INT	intentional
FAM	familiar	INTER	intermediate
FAR	far	INTR	intransitive
FRST	firsthand evidentiality	INTRG	interrogative

INV	inverse	PASS	passive
IPD	impeditive	p.c.	personal communication
IPFV	imperfective	PERI	peripheral (participant)
IPS	impersonal	PFV	perfective
IRR	irrealis	PL	plural
ITE	iterative	PLPF	pluperfect
ITG	intangible (dem.)	PNC	punctual
IVN	instrumental verbal noun	POL	polite
JUS	jussive	POSS	possessive
LCT	locutor person marker	POSTP	postposition
LIG	ligature	POT	potential
LIM	limitative	PR	pronoun
LK	linker	PRD	predicative marker
L	long form	PRI	progressive intransitive marker
LM	long form male speech	PROH	prohibitive
LOAF	low affectedness Aktionsart	PROG	progressive aspect
LOC	locative	PROP	propriative
LPM	locutor person marker	PROX	proximate, proximal
M	masculine gender	PRPS	progressive persistent
MED	medial deictic aspect	PRS	present tense
MNR	manner	PRSN	presential
MOT	motion	PST	past tense
N	neutre gender	PSTP	postposition
N	non- (as in NPST =non-past NF=non-feminine)	PTC	particle
NCL	noun classifier	PTCP	participle
NEG	negative, negation	PRP	purposive, purpose
NMLZ	nominalizer	PVN	plain verbal noun
NOM	nominal	QUOT	quotative
NR	subject nominalizer	R	respectful
NSAP	non-speech act participant voice	RE	restorative
NTR	neutral form	REA	realis
O	O-participant; O-marking	REAS	reassurance
OBJ	object	REC	recent
OBQ	oblique	REDUP	reduplication
		REF	referential

REFL	reflexive	TAME	Tense, Modality, Aspect, Evidentiality
REI	reiterative	TEL	telic
REL	relative	TEMP	temporal
RELN	relational	TERM	terminative
REM	remote	THEM	thematic vowel
REP	repetitive, repetition	THI	thither
REPO	reportative, reported	THRD	thirdhand
RES	resultative	TMPR	temporarily
S	S participant	TOP	topic(al)
S _A	S _A participant, S _A class marker	TOT	totalizer
S _O	S _O participant	TR	transitive
SA	South America	UNPOS	unpossessed noun
SAIL(s)	South American Indigenous Language(s)	VAL	validational
SAP	speech act participant voice	VBLZ	verbalizer
SBJ	subject	VIS	visual evidentiality
SCND	secondhand	VLC	locative verbalization
SG	singular	VM	verbal stem marker
SI	specific inferred	VN	verbal noun
SIT	situational	VOC	vocative
SM	short form male speech	VPL	verbal plural
SML	similative	VS	verbal suffix
SOC	sociative	VSB	visible
SP	speaker		
SPC	specifier		
SR	switch reference		
SS	same subject		
SSSS	simultaneous event, same subject, S-orientation		
STA	stative		
STV	stativizer		
SU	superessive relation		
SUB	subordinator		
SUP	supine		
SV	serial verb marker		

1. INTRODUCTION

This study presents an analysis of morpho-syntactic Tense, Aspect, Modality, and Evidentiality marking (henceforth: TAME) in a sample of 63 South American indigenous languages (henceforth: SAILS) with regard to typological, geographical, and genealogical distributions.¹ This chapter introduces the study and serves to embed it in the current research on South American linguistics. An overview of the structure of the book is given at the end of this chapter. The main goals of this thesis are twofold:

- (i) to present a typological profile of morpho-syntactic Tense, Aspect, Modality, and Evidentiality marking in a sample of 63 South American indigenous languages,
- (ii) to uncover genealogical and geographical relationships of SAILS according to the TAME profile with special focus on language contact.

This is the first comprehensive and comparative study of TAME in South American languages in a broad sample and it will hopefully serve as a basis for both detailed studies of TAME categories in SAILS as well as those of a global typological nature. A questionnaire was designed for Tense, Aspect, Modality, and Evidentiality, and the data of 63 languages were entered in full, with various additional languages for specific subtopics (see section 2.5.2).

Until recently, there was a discrepancy between the enormous range of SAILS and their deficient documentation, but in the last decades the increasing number of high-quality descriptions has made it possible to conduct large scale research on SAILS. A number of research questions have arisen, most prominently with regard to the high genealogical diversity and the spread of certain language families, but also to the typological characteristics and how they contrast with other parts of the world. For instance, the most recent count by Hammarström (2009, appendix) presents 111 language families (or isolates) in South America. For comparison, there are only 33 families in Eurasia (*ibid.*).

The nature of the topic, the methodology, and the language sample present their own sets of problems, and it is important to list them here. First, the categories of Tense, Aspect, Modality, and Evidentiality stand out by their closely connected semantic relationships and the difficulties to demarcate the categories from each other. Second, the language sample is a convenience sample and therefore any extrapolation on the basis of the results must be taken with caution. Third, the very fact that there are at least 63 different sources with different opinions of terminology by the authors that have I have to reconcile with my own chosen definitions. Fourth, even though mostly well documented languages were chosen, the available grammatical descriptions are often not sufficiently precise.

¹ I acknowledge the financial support from the ERC project “Traces of Contact”.

It became apparent quite early in the study that a full analysis of TAME systems in the sample was not feasible in the defined time frame. Because I was inspired by recent work on grammaticalization (cf. Bybee et al. 1994) and its relevance for borrowing (cf. Heine & Kuteva 2005) I decided to limit the study to morpho-syntactic marking.

Recently, attempts have been made to infer population pathways of the South American continent by relating the distribution of languages to archaeological, anthropological, and data from similar research fields (cf. Eriksen 2011, Hornborg & Hill 2011) by applying both old and new research techniques. For example, the study of phylogenetics has been adopted from biology to reveal relationships of languages that have yet to be classified, and to confirm language families (cf. Dunn et al. 2008). Although it is disputed whether phylolinguistics yields significant results this is just one example of the new and exciting possibilities to increase our knowledge of SAILS where traditional methods fail.

Each source was carefully checked although in the case of e.g. bad data that was not always possible. In several instances, my analysis deviates from the sources. That, of course, risks bringing the wrath of the individual language specialists upon me. As Stassen (1997, preface) aptly puts it: “Thus, the primary aim of my research project has been theoretical and universal in nature, but it is evident that it relies heavily on the descriptive work done by specialists on singular languages or language groups. [...] As is always the case with go-betweens, this puts typologists in a somewhat uncomfortable position, in which one runs the risk of being shot at from both sides”. I welcome any shots from any side as opportunity to expand on my scholarly education.

This book is structured as follows: First, the framework and methodology are explained in chapter 2, together with presenting the language sample and the structure of the questionnaire. Chapters 3 to 6 constitute the main body of the thesis and investigate in detail the typological patterns and geographical and genealogical distributions. The order is Tense (chapter 3), Aspect (chapter 4), Modality (chapter 5), and Evidentiality (chapter 6). Each chapter begins with a definition of the respective categories. Chapter 7 aims to consolidate the results from the previous chapters and examine their categories in a comprehensive manner, with special focus on geographical and genealogical distributions, TAME prominence, and temporal stability of TAME features.

2. TERMINOLOGY, FRAMEWORK, AND TOOLS

2.1 INTRODUCTION

The present study focuses on the occurrence of morpho-syntactic TAME markers in 63 selected SAILS. It aims to illustrate the typological TAME landscape of SAILS in a comparative perspective and to serve as a starting point for historical linguistic studies as well as modeling language contact scenarios. This chapter presents the methodological foundations of the study. First, the specific terminology of ‘marking’ and what kinds of markers occur in the sample are discussed in section 2.2. Section 2.3 presents the frameworks central to this study: the prototype theory, the canonical approach, and the dominance parameter. Because this study focuses on grammaticalized markers, grammaticalization is briefly discussed in section 2.3.3 together with its related notions of obligatoriness (2.3.4) and systematicity (2.3.5). Section 2.4 presents the *modus operandi* of this study, based on section 2.3, and offers some general remarks about the levels of description in this study and what is meant by ‘morpho-syntactic’. The language sample and the structure of the questionnaire are presented in sections 2.5. The definitions of the individual TAME categories are given in the respective sections in chapters 3 to 6. When possible, examples are given to illustrate the point in question. Examples usually consist of three lines: morpheme by morpheme original in the first line, glossing in the second line, and the English translation in the third line. In cases where the source translation is not in English, an English one is supplied. In some cases, the original examples do not have glossing or morpheme separations; it was not attempted to substitute these. In general, those examples were avoided (but see e.g. example (5.86)). It was attempted to keep as true to the examples in the sources as possible, but with some alterations: glossings were unified and simplified in accordance with the Leipzig glossing rules, and in some cases existing glosses were replaced for purpose of illustration. For example, in (6.21a) the particle *topə* is glossed ‘vis’ instead of the original gloss ‘be.seen’ to signal that *topə* belongs to VISUAL Evidentiality as defined in this study. In cases where the original gloss seems to be a mistake, it was replaced by the most obvious candidate. For example, in example (6.1), the glossing ‘AU’ was replaced by ‘AUX’, because ‘AU’ does not occur in the list of abbreviations, but ‘AUX’ does.

2.2 TERMINOLOGY

2.2.1 Marking

Haspelmath (2006, 27) argues that linguists should refrain from using ‘unmarked, marked’ terminology, because these terms “developed a multiplicity of sometimes widely diverging senses”, and he suggests to replace them by ‘overtly coded, uncoded/ zero coded’ (ibid, 30). I agree that there has been confusion about these terms, but would like to add that ‘zero coded’ is not necessarily the same as ‘uncoded’ and that one may use both ‘marking’ and ‘coding’ as long as they are clearly defined. In this thesis, both sets of terms are used interchangeably denoting the following concepts: ‘marking, coding’ means that a category

has a visible marker containing phonological/ morphological material. In that sense, a category can only be marked or unmarked, but not less or more marked. An example where degrees of marking are discussed is Muysken (1981) who has attempted to show degrees of Tense markedness. Counting the distances between point of reference, point of speech, and event moment, Muysken concludes that the higher the distance, the more marked the category in case. For example, in the simple present all three points overlap, and the distance is zero. In the simple past, the points of reference and speech overlap, but the event moment is removed, so the distance is one. Therefore, the simple past was assumed to be more marked than the simple present (ibid. 190–191) (for a discussion of the terms point of reference, speech, and moment of utterance, see section 3.2.1). The present thesis does not attempt to rank categories according to markedness, but rather sorts them into marked vs. unmarked. In Karo, a clause can be marked with the PAST particle *co* or be unmarked for PAST but still have past time reference. Additionally, PAST can be indicated by an adverb ('yesterday'), and be both marked by a particle and indicated by an adverb in the same clause:

(2.4) Karo (Tupían; Gabas Jr. 1999, 175, 98, 169, 175)

- (a) *púŋ o=?e-t co*
 shoot 1SG=AUX-IND PST
 'I shot.'
- (b) *o=ket-t cú=tem*
 1SG=sleep-IND big=ADVZ
 'I slept a lot.'
- (c) *mēganape òn ameko top-ap matet*
 here 1SG jaguar see-IND yesterday
 'Here I saw a/ the jaguar yesterday.'
- (d) *iyòm ket-t co matet*
 father sleep-IND PST yesterday
 'Father slept yesterday.'

In my approach, (2.4a) and (2.4d) are morpho-syntactically marked for PAST, but (2.4b) and (2.4c) are not. (2.4d) is also not more marked because it has both a clitic and an adverb.

Periphrastic constructions and adverbs are not taken into account in this thesis except when explicitly stated otherwise. This thesis focuses on grammaticalized TAME marking, although other means of expressing TAME are discussed in appropriate places. The opposites of 'marked, coded' are 'unmarked, uncoded', when a category lacks a visible marker.

Furthermore, a word about the terms 'unmarked' and 'zero-marker' is in order. In this study, both unmarked and zero-marked categories are treated the same, although they do not denote the same concepts. I follow Bybee (1994) who argues that a zero-marker is a marker with semantic content, but without overt realization in the form of a morpheme. It

is therefore the absence of an (obligatory) overt marker that expresses the existence of another, related meaning: “[a]s an overt marker becomes more frequent, the hearer can infer that its absence is intentional and meaningful, leading to the development of zero-grams¹” (ibid. 252). Bybee (ibid. 240) claims that several criteria must apply to a marker to be a real zero-marker, including the increase of frequency of a marker that, additionally, is obligatory and whose absence then can be inferred to be meaningful. A category that is unmarked, on the other hand, neither has an overt marker nor does it lead the hearer to infer a specific meaning. For example, in many languages the PERFECTIVE has become zero-marked after the IMPERFECTIVE became obligatory (ibid. 250). Unfortunately, due to different uses of the term ‘zero-marker’ in the sources and the temporal limitations of the project it was not possible to distinguish between zero-marked and unmarked categories in this study. It is especially difficult to infer obligatoriness and frequency of markers in reference grammars. For that reason, in this thesis an unmarked category may well be zero-marked but is undetected as such. However, TAME zero-marking in SAILS may be worth investigating in the future.

2.2.2 Cumulative morphemes

In SAILS, there are often cases where one marker expresses two or more TAME categories simultaneously in an indivisible unit. Bauer (2003, 337) calls this ‘portmanteau’: “A portmanteau morph is a morph which realizes ... more than one morpheme”. Bickel & Nichols (2007, 188) refer to portmanteau morphemes (or portmanteau formatives) as combining features that also exist independently. For example, French *du* is a portmanteau of *de* ‘of’ and *le* ‘the’. The present study uses Bickel & Nichols’ (2007) term ‘cumulative’ for a morpheme that combines two (or more) features, but without coexisting independent forms (as would be the case with portmanteau morphemes). A prototypical cumulative marker here combines two or more meanings which always apply at the same time. For example, in Tiriyo, the suffix *-ja* marks PRESENT Tense and IMPERFECTIVE Aspect, and in a clause this marker expresses both. That means that if one applies, the other obligatorily applies as well.

(2.5) Tiriyo (Cariban; Meira 1999, 299)

fevereiro po tarëno-ton eperu pëë-ja-n
 February LOC Tiriyo-COL fruit gather-PRS.IPFV-DUB
 ‘In February, the Tiriyo gather fruits.’

Cumulative markers (henceforth also called cumulatives) can be cross-categorial, but never intra-categorial: a marker may have two meanings from either Tense, Aspect, Modality, or Evidentiality, for example PAST and PERFECT, but never from the same super-category, e.g.

¹ Short for “grammatical morpheme” (Bybee et al. 1994, 2).

PAST and FUTURE. When a marker exhibits two (or more) meanings from the same category, such as PRESENT and PAST, this marker is defined as not cumulative, but ambiguous (see the following section).

2.2.3 Ambiguity

In TAME, morphemes similar to cumulatives also occur: ambiguous morphemes. A cumulative morpheme always comprises several meanings at the same time, but an ambiguous morpheme, although having several meanings, only has one meaning in a specific context. A typical example is a binary Tense distinction, e.g. of the type FUTURE/ NONFUTURE where the NONFUTURE marker points to either a PRESENT or PAST interpretation, i.e. is ambiguous between PRESENT and PAST. This is not a cumulative morpheme, because in a clause not both, but only one meaning applies. For example, in Hixkaryana the marker *-yaha* applies to NONPAST, i.e. one marker codes for either PRESENT or FUTURE (and it can also be interpreted as ‘universal’) (Derbyshire 1979, 138). Which Tense is actually applicable in the clause has to be established by context, but it can never be both at the same time. Ambiguity is typically intracategorical, i.e. usually occurs between features within one category, in opposition to cumulatives which are always cross-categorical.

(2.6) Hixkaryana (Cariban; Derbyshire 1979, 138)

namryekyaha

‘he is going hunting (now)’ or ‘he will go hunting (sometime soon)’, or ‘he hunts’
(i.e. he is a man who hunts).

The questionnaire does not make a distinction between cumulative and ambiguous markers in the sense that both get a positive entry in all of the categories they encode. Specifications are made in the comment section.

Cumulative and ambiguous morphemes are interesting from a diachronic point of view, since the presence of such morphemes in a language points towards a high degree of grammaticalization. This study will discuss the degree of exponence (i.e. categories that form cumulatives) where relevant.

To outline the problematic status of TAME regarding ambiguity I now discuss a few complicated examples. For example, a PERFECTIVE marker can also express PAST Tense, an IMPERFECTIVE marker can have HABITUAL functions, a FUTURE Tense marker can be used to express INTENTIONAL, and a direct Evidential may have Epistemic undertones. It is impossible to establish clear-cut borders and instead one should regard the categories as clusters that share traits. The clusters have prototypical meanings that are central to a special feature, e.g. PAST Tense refers to a point in time previous to the point of speech and/or reference. There are features that share this definition, e.g. the PERFECTIVE can have PAST Tense meaning, but is distinguishable by stating that this is not the core meaning of PERFECTIVE, just a side effect, or secondary meaning. According to Dahl (1985), the core meaning is ‘focusing’, and the grey area in between two features ‘imprecise’. He claims that the focus of

a category equals a prototype and assigns a major status to the “concept of a dominant parameter”, i.e. a feature may have several traits, but only one is dominant and therefore the core meaning. This approach is in a position to account more or less satisfactorily for TAME relationships and it will be followed here (see following sections).

There is a difference between cumulative and ambiguous morphemes and morphemes that seemingly have two functions but under close scrutiny just have one dominant meaning and another one resulting from it (secondary meaning). Unfortunately, it is not always possible to discern which is dominant and which is secondary. In case one is not sure to which category a certain marker belongs (e.g. PAST Tense or PERFECTIVE Aspect), one should establish the core meaning. A marker that always has past time reference, but only sometimes marks present relevance of a past event is predominantly PAST Tense marking with a secondary ANTERIOR meaning. In case the marker seems to always mark both, refer to the discussion about cumulative morphemes above. In that way, one can make a list of values of a marker, e.g. a marker may have the following values: [+S<E], [+intentional], [+uncertainty], but not any of the other definitions. Thus, the relevant marker marks FUTURE, INTENTION, and UNCERTAINTY. Ideally, a source indicates the status of these meanings, but that is not always the case. The biggest problem with this approach is that the specific language source necessarily has to have a fair amount of examples. There is also the risk of being misguided by a translation.

2.3 CONCEPTUAL FRAMEWORK

2.3.1 Prototype theory and the canonical approach

The semantic categories of Tense, Aspect, Modality, and Evidentiality are challenging as they are not clear-cut. In order to define language-specific TAME markers, but also in order to establish cross-linguistically valid concepts, I discuss two typological approaches that are important to structure and distinguish the categories: (i) the prototype theory, which has already been applied to Tense and Aspect by Dahl (1985), and (ii) the canonical approach which was introduced by Corbett (2003) and applied to the phenomenon of agreement.² The following section illustrates how these approaches apply to the study of TAME.

Originally developed in the 1970s by the cognitive psychologist Eleanor Rosch and colleagues, prototype theory was first applied to Tense and Aspect in a cross-linguistic study by Dahl in 1985. The most basic problem in defining and distinguishing TAME features is their fuzzy boundaries, and prototype theory (also called prototype semantic theory), adopted to the special needs of TA, offers the, in my opinion, best suited approach to keep these categories apart. The arguments of the standard prototype theory are that, firstly, there are categories and that, secondly, some specimen of these categories can be arranged into being ‘better’ specimen than others, e.g. an *apple* is a better specimen of the category *fruit* than an *olive* is (see Kleiber 1998 for examples). A prototype is then the best specimen of a category by vote of the majority: “[t]he term ‘prototype’ as best specimen or rather

² For the canonical approach being applied to the Miraña agreement system see Seifart (2005).

example, best representative or central element of a category [...]. The prototype is the specimen that is deemed most fitting by the speakers.” (Translation N.M.) (Kleiber 1998, 31).³ A prototype is not an individual specimen (e.g. for *bird* the *parrot of my neighbor*) but a subcategory (*parrot*). Rosch et al. (1976) established the prototypes in experiments by asking people which specimen of, for example, a series of chairs best represented the category chair. They then formed a scale of chairs, the highest of which was the prototype. This cannot be done for TAME, however, so that we have to establish the prototypes for TAME categories in a different way. The advantage of the prototype theory is that it is flexible, i.e. allows for less prototypical members of a category. The disadvantage is that TAME prototypes cannot be established by vote of majority and that not even two scholars may actually agree on them. A further general disadvantage of the prototype theory is that it is probably culture-specific. Defining TAME categories for language comparison requires moving from language-internal features to more abstract concepts, which the prototype theory cannot give us. It is not sufficient to choose e.g. English FUTURE as prototype and compare all other language specific FUTURES against it; an ideal concept of FUTURE is needed. This is the topic of the next paragraph.

Another typological approach, which to my knowledge has not been applied to TAME before, is the canonical approach by Corbett (2003, 2007). Whereas the prototype theory works with a bottom-up approach, i.e. draws definitions from empirical data, the canonical approach combines a top-down and bottom-up approach in that it proposes a definition and maps it against actual data. It has a core definition that is the result of “tak[ing] definitions to their logical endpoint” and which allows “to build theoretical spaces of possibilities” (Corbett 2007, 9). So instead of choosing a best exemplar prototype (e.g. English FUTURE) the canon provides an ideal concept (the point of event is after the point of speech) that is then mapped against language-specific instances; this is what Dahl means with primary meaning. Both prototype theory and canonical approach allow for gradation. In a sense, a canonical instance is the same as Dahl’s prototype. Dahl’s prototypes deviate from the original prototypes in that they are not the best example of a group, but rather an ideal, which is meant by a canonical instance. Thus, Dahl abstracted the original prototype theory and enhanced the meaning of a prototype to denote an ideal, which has later been labeled ‘canonical instance’ by Corbett.

2.3.2 The dominance parameter

A TAME marker rarely has only one meaning, but in order to be able to sort it into a category one needs to decide which category it belongs to by establishing its primary meaning in terms of dominance. In a language X with a marker combining FUTURE and e.g. UNCERTAINTY, one has to look at a set of examples to establish which one is dominant. When all examples have FUTURE meaning but only sometimes a Modal meaning as well, FUTURE is

³ “der Begriff des Prototyps als bestes Exemplar bzw. Beispiel, bester Vertreter oder zentrales Element einer Kategorie [...]. Der Prototyp ist das Exemplar, das von den Sprechern als bestes anerkannt wird”.

the dominant meaning, and the marker is therefore a Tense marker. Dahl (1985, 9) calls this a “concept of dominant parameter” which “is often relevant in the description of TAME categories. For example, the category Perfective will usually be interpreted as ‘perfective’ and ‘past’, although with the first feature clearly dominant”. In relation to the canonical approach, this means that the canonical meaning is the dominant one. It is of course possible that markers have several meanings that always occur in all instances, i.e. they are cumulative morphemes (c.f. section 2.2.2).

2.3.3 Grammaticalization

The inherent semantic properties of TAME categories suggest that TAME features are part of a continuum. They are strongly related to each other, to the point of overlapping. Within this continuum, markers are prone to shift meaning, some more easily than others, and also to shift form and position in the clause. It is this conceptual linkage between the categories that facilitates grammaticalization processes between TAME categories. Shift of meaning is thus a prominent characteristic of grammaticalization or ‘grammaticization’ (cf. Bybee et al. 1994, Heine & Kuteva 2005, Narrog & Heine 2011, to name but a few). The following paragraphs briefly introduce the concept of grammaticalization and why it is so important in our understanding of TAME categories.

Grammaticalization is a process in which a certain lexeme undergoes changes in meaning and form, and evolves to the status of grammatical element. Changes in meaning include a generalization of the original semantic content, or reduction (erosion, bleaching, narrowing etc.). Changes in form include reduction of morpho-phonological material, i.e. shortening or condensing. The change of meaning and form subsequently leads to a change of word class as well, i.e. a formerly independent lexeme is found in a much more restricted environment such as the affix slot. The various changes a lexeme undergoes to become more or less grammaticalized are commonly referred to as grammaticalization path or cline. This is a very general definition of grammaticalization, based on Bybee et al. (1994) who investigated Tense, Aspect, and Modality regarding grammaticalization paths in a global sample of 76 languages. According to them, a grammatical marker exists on a special point along a grammaticalization path and can only satisfactorily be described diachronically in a way that allows for cross-linguistic comparison: “[s]ynchronic universals have been elusive in the area of grammatical meaning because at any given time a gram’s uses stretch over a sequence of links in the grammaticalization chain. Only the diachronic perspective can reveal how these uses are related and how a given gram compares to similar grams in other languages” (ibid. 281). In other words, for a marker to be comparable one should know where it came from and at which point of grammaticalization it exists now, both in meaning and in form. This is, of course, hardly feasible for SAILs, because for most of the languages we do not have sufficient historical data for diachronic analyses. However, Bybee et al. (ibid.) provide us with a number of universal grammaticalization paths that enable us to infer probable sources of specific markers. In addition, a general interest in grammaticalization has also resulted in a number of studies of particular

phenomena in SAILS (e.g. Epps 2008 about the grammaticalization of a FUTURE marker from the lexeme for ‘wood’ in Hup). Although inferring grammaticalization paths of TAME markers in SAILS is not a major research topic of this study, they are of considerable help when sorting and comparing markers.

Grammaticalization does not only contribute to our understanding of TAME markers and their language-internal development, it has also been demonstrated that it can be triggered by contact and thus reveal contact in contiguous regions with unrelated languages. Thus, contact of a language with a certain grammaticalized marker can trigger grammaticalization of a morpheme in an adjacent language, as argued, among others, by Heine & Kuteva (2005). A region that features several instances of transfer by grammaticalization of unrelated languages is called ‘grammaticalization area’: [b]y grammaticalization area [...] we understand a group of geographically contiguous languages that have undergone the same grammaticalization process as a result of language contact” (ibid. 182). Heine & Kuteva propose several grammaticalization areas world-wide, and one specifically in SA. In the Vaupés, contact-induced grammaticalization is known to have happened between Tucanoan and Tariana (Arawakan). For example, the category of a SECONDHAND IMPERATIVE (“Do this because someone else told you!”) was transferred from Tucanoan to Tariana. Tariana did not originally have a SECONDHAND IMPERATIVE and instead extended the meaning of the SECONDHAND marker *-pida* to IMPERATIVE (ibid. 214). It is beyond the range of this study to propose grammaticalization areas in SA, but contact-induced grammaticalization may explain the distribution of certain TAME markers, just like language internal grammaticalization does.

2.3.4 Obligatoriness

Grammaticalization, especially that which eventually leads to a zero-marker, is most likely when one or more meanings are obligatorily expressed in a specific language (Bybee 1994, 252). In general, any category is obligatory when a clause unmarked for that category meaning is ungrammatical. That meaning is usually tied to a specific marker or a paradigm of markers. For Bybee (1985, 202), a category is obligatory when it is compulsory in every finite clause regardless of whether its meaning is redundant. Unfortunately, languages are not that persistent with regard to obligatoriness. For instance, in Tariana grammatical Tense marking is obligatory but can be dropped, for instance, in narratives when the temporal frame has been established (Aikhenvald 2003, 289).

Apparently, there is a difference between obligatoriness of marker and obligatoriness of meaning. When the meaning is expressed solely by one marker, then both coincide, but when the meaning is possibly expressed by alternative means, they overlap only partly. That explains why it is possible to drop a Tense marker in narratives or subordinate clauses, because the meaning is still there, although the overt marker is not.

In this study, it has not been possible to systematically differentiate between the obligatory and optional marking of a category. This is due to the fact that many sources simply do not refer to obligatoriness, and that when they do, they are not entirely

consistent in their definitions. This is one of the factors that led to the exclusion of zero-markers as well, because obligatoriness of one category is vital for the grammaticalization of another category into a zero-marker. Therefore, this study includes markers whose obligatoriness is unknown.

2.3.5 Systematicity

TAME markers, especially highly grammaticalized ones, are strongly embedded in grammatical systems, i.e. are part of complex paradigms. A category is taken into account in this study regardless of whether it is represented by markers in all instances in a paradigm or only partially. Paradigms can be defective. For example, Huallaga Quechua has a FUTURE paradigm for person and number, but the forms for the second person are homophonous with those of the PRESENT (Weber 1989, 100). In Tiriyo, the non-collective PRESENT IMPERFECTIVE and FUTURE IMPERFECTIVE have a CERTAINTY/ DUBITATIVE distinction that is neither present in the collective forms nor in the PRESENT PERFECTIVE, FUTURE PERFECTIVE, or in the PAST Tenses (Meira 1999, 295). As languages can have only partial obligatoriness, they also can have partial paradigms. Nevertheless, markers of defective paradigms are still taken into account here. For Tiriyo, for example, that means that Tiriyo is coded to have marked Epistemic Modality, even though only in PRESENT IMPERFECTIVE and FUTURE IMPERFECTIVE.

2.4 THE MODUS OPERANDI OF THIS STUDY

This section explains the fundamental framework of this study which was partially introduced in section 2.3.

In this study, the definitions given in the respective sections are the ideal, i.e. canonical/ prototypical, meanings which were inferred from previous studies. Language-specific markers may more or less canonically/ prototypically represent these meanings. For example, FUTURE is generally accepted to be the grammaticalized expression of time constituting that the point of speech is before the point of event. Accordingly, this is the definition of FUTURE used in this study (cf. section 3.2.1). But markers expressing FUTURE are not necessarily restricted to this ideal definition. For example, a FUTURE marker may also express DUBITATIVE. In that case, one needs to establish which meaning is dominant, or whether these meanings are equally dominant (i.e. a cumulative morpheme).

Dahl (1985, 33) argues that three levels of description have to be present in a successful typological study of TAME: one for universal categories, i.e. PAST, followed by the level of language-specific categories (e.g. PAST in English). The third level denotes the dimensions of conceptual space, i.e. the 'quarks' that make up the space of a prototype. For example, the universal category PAST is realized in the language-specific category PAST Tense in English, and the meaning of English PAST is past time. But past time is not restricted to PAST Tense and may in some languages also be a part of PERFECTIVE Aspect. Naturally, past time reference and PAST Tense coincide, as the semantic content of PAST Tense is past time, but a category is not restricted to have only one quark. FUTURE Tense often includes not only future time reference but also Modal meanings like IMPERATIVE, planning, desire etc., depending on the

individual language. The quarks are the components that need to be semantically defined, as through them the higher categories get their meaning.

This thesis works with two basic distinctions: supercategories and their respective subcategories. Supercategories are indicated by an initial capital, e.g. Tense, Modality, Aspect, and Evidentiality to distinguish them from homophones such as tense and mood, as in “I walked into a bar and the mood was quite tense”. Subcategories are formatted in small caps, regardless of whether they also represent subcategories of subcategories, e.g. IMPERFECTIVE and CONTINUATIVE are not formally hierarchically distinguished although CONTINUATIVE is a subcategory of IMPERFECTIVE (cf. section 4.1). This is done in order to establish visible differences between terms like PAST and past, the former being the linguistic category and the latter referring to a general concept of time, not being restricted to the linguistic definition. Categories that are not dealt with, and thus not defined, in this study receive no extra marking. For example, interrogative, negation, indicative, etc. are recognizable as not being part of the questionnaire. The supercategories are assumed to be language universals for the purpose of cross-linguistic comparison, although this study does not claim that they are.

Figure 2.1: Levels of description

<u>Dahl 1985</u>	
Universal category	e.g. PAST
Language specific category	e.g. PAST in English
Quarks	e.g. past time reference
<u>This study</u>	
Supercategory (universal)	Tense, Aspect, Modality, Evidentiality
Subcategory	e.g. PAST, REALIS, PERFECTIVE, FIRSHAND
Definition	e.g. past time reference

As pointed out by Haspelmath (2010), individual language descriptions necessarily work with descriptive sets applicable to the respective language, which means that they are not comparable cross-linguistically, as no particular category has an exact one-to-one correspondence to a category of another language. Instead, one needs to work with comparative concepts that are “specifically designed for the purpose of comparison that are independent of descriptive categories” (Haspelmath 2010, 664). The problem of language-particular categories that denote different concepts from language to language is deeply rooted within TAME, not only because of different language internal processes, but also because of the tendency of researchers to apply terminology that is not sufficiently defined. Thus, two language-specific IRREALIS markers may well be different in all but terminology, but that should not pose a problem as long as both uses are defined. Some authors of descriptive grammars first often do not make a distinction between language particular and universal categories, and second do not always define their terminology. This is aggravated

by the fact that TAME categories are highly interrelated and without definition one is unable to distinguish features and ends up comparing two categories that are not the same at all, simply because both have been labeled the same.⁴

Haspelmath's comparative concepts approach offers a solution for the way to find linguistic universals in comparing not linguistic features, but concepts. A modified version of comparative concepts can be found in this thesis. The bases for cross-linguistic comparison are the categories, which are defined by prototypical/ canonical meanings (see below). It should be borne in mind that e.g. the category PAST does not necessarily fit one-to-one to language specific PAST markers.

This study focuses on TAME categories that are marked morpho-syntactically. Due to the limits of the project periphrastic constructions, adverbs, full verbs etc. are not taken into account, although they are probably not less interesting with regard to language contact, language change, grammaticalization, and typology. It is hoped that a complementing study of non-morpho-syntactic TAME of the same sample will be conducted in the future. Morpho-syntactic marking here includes affixes, clitics, particles, auxiliaries, and stem change/ suppletion (and for some questions, also repetition/reduplication). These are partially distinguished by their ability to occur as bound or free morphemes and their ability to take morphology: affixes and clitics cannot occur on their own, but have to be attached to a host, whereas particles and auxiliaries are free morphemes, although partly restricted to certain positions in the clause (language-dependent). A particle cannot take affixes or clitics, whereas auxiliaries are able to receive verbal inflectional morphemes. The boundaries are fluid, which leads to classification problems. For example, in Baure (Arawakan), several adverbs and particles are homophonous and obviously the particles derive from the adverbs, e.g. the PERFECT particle "has been derived from the adverb *ver* 'already', even though the adverbial use is hardly distinguishable from the particle use" (Danielsen 2007, 272). In general, the terms 'particle' and 'adverb' are often used interchangeably ('particle' is also infamous for being used for every marker that could not be sorted into any other word class) and one has to rely on the original source information.

A TAME marker always has scope over the verb phrase, regardless of its position in the clause. For example, a marker that attaches to nouns and has no scope over the verbal action of the clause is not a Tense marker. But when that same marker also has verbal scope, then it is a Tense marker.⁵ For example, in Apurinã the FUTURE marker *-ko* can occur on non-verbal bases as well as on the verb, but it always has scope over the verbal action (Facundes 2000, 410). In Mamaindê there is a set of temporal markers that exclusively occur on nouns (i.e. nominal Tense markers), in addition to a distinct set of verbal Tense markers (Eberhard 2009, 343). Both languages are coded as having Tense markers, but the Mamaindê

⁴ This particularly happens with the terms 'perfect' and 'perfective' (cf. section 4.7).

⁵ For a discussion of the nominal tense see Nordlinger & Sadler (2004, 2008), Tonhauser (2007), and (Muysken 2008).

nominal Tense markers are not taken into account in this study, because they do not have verbal scope.

- (2.1) Apurinã (Facundes 2000, 410, 514)
- (a) *kopiti-ka-ra-ko* *o-kama* nominal base
 bucket-PRD-FOC-FUT 3F-make
 ‘(It’s) the bucket that she’ll make.’
- (b) *iye* *Ø-oka-pe-no-na-ko* verbal base
 PTC 3M-kill-PFV-1SG.OBJ-3PL-FUT
 ‘So, they will kill me.’
- (2.2) Mamaindê (Nambikwaran; Eberhard 2009, 343, 429)
- (a) *ta-walek^han-leʔi-tu* nominal Tense
 POSS1-chief-PST-FNS
 ‘my past chief’
- (b) *jalakwatun-tu* *ãn-let-Ø-nãn-wa* verbal Tense
 howler.monkey-FNS shoot-I.PST-SBJ3-PST-DECL
 ‘He shot a howler monkey (in intermediate past time).’

Markers can occur on auxiliaries and also be auxiliaries themselves: in Shipibo-Konibo the main verbs *peo-* ‘begin’, *keyo-* ‘finish’ and *jene-* ‘stop’ can also function as auxiliaries with Aspectual meaning (Valenzuela 2003, 319). In Dâw, the PERFECTIVE marker is an auxiliary derived from the verb *jūt* ‘to kill’ (Martins 2004, 290):

- (2.3) Dâw (Nadahup; Martins 2004, 290)
- (a) *m^ʔɛʔ* *jaʃ* *pɣʔ* ***jūt*** *mɛt* verb
 one time/journey grandparent kill cutia
 ‘Once the grandmother killed an agouti.’ (orig. ‘Uma vez a avó matou uma cutia.’)
- (b) *tih* *pow* ***jūt*** *bɛ-duh* auxiliary
 3SG to.split PFV plant-firewood
 ‘He split firewood.’ (orig. ‘Ele rachou lenha.’)

To summarize, the present study takes into account morpho-syntactic TAME marker. The definitions of the TAME categories in this chapter are ideals/ prototypes/ canonical instances based on previous studies. A marker must have the prototypical meaning as the dominant one. In case there are two or more dominant meanings it is either a cumulative or ambiguous marker. It does not matter whether a marker is cumulative or ambiguous, because in both instances the category is coded as marked. For further information about coding strategies see section 2.5.2.

2.5 TOOLS

2.5.1 The language sample

The language sample used in this study was designed to balance geographical and genealogical diversity of SAILS with a focus on certain families and regions, taking the availability of good descriptions into account. The basis for the analysis is a questionnaire, which is the topic of the following section. In this section I present the language sample and how it was compiled.

The basis of this particular sample is a list of 60 languages which was compiled for the ERC project “Traces of Contact (ToC)”. One of the main goals of the project is to integrate the growing body of descriptive work which has become available in the last decades. For this study, three additional languages were chosen in order to expand the coverage of major language families: Quechuan (Cuzco Quechua) and Tupían (Nheengatú and Sateré-Mawé) so as to allow for more detailed genealogical analysis. The original ToC language sample was compiled based on three principles: first, to profile the known major language families as well as smaller ones and isolates/ unclassified languages (genealogical diversity); second, to present a broad geographical coverage (geographical diversity); typological diversity; and fourth, availability of adequate descriptions (cf. Krasnoukhova 2012, 7-8). The additional languages are also chosen according to these principles, but with a small specification: the available descriptions should include satisfactory information about TAME. An ideal source should adhere to all of the following principles:

1. Quality: preferably descriptive grammars
2. Quantity: detailed discussions and presentation of examples with glossing, also annotated text examples
3. Contact with language specialist(s)
4. Hitherto under- or undescribed languages

The sources in the sample include a variety of the characteristics above. For example, there are several reference grammars excellent in both quality and quantity, e.g. Mapuche (Smeets 2008), Jarawara (Dixon 2004), or Nheengatú (Da Cruz 2011), which are complemented in some cases with personal communication with the authors (e.g. Kwaza, Mosestén, Tiriyo, Yurakaré). There are also a few cases in which a source includes only articles, but where contact with language specialists could close gaps of information (e.g. Leko, Itonama). If contact could be established with specialists, they were usually very approachable and helpful with particular questions, but often enough the researcher could not be located or did not respond. Unfortunately, although it was expected that a comprehensive grammar of Panare would soon be published, this did not happen during the course of this study.⁶ It was replaced by Gildea (1992, 1998, 2012). A detailed list of the languages in the sample and the sources can be found in table 2.1. The following language

⁶ It was finally published at the end of 2012 (Payne & Payne 2012).

families are represented in the sample (the number in brackets refers to the number of members in the sample):

Arawakan (5), Arawan (1), Araucanian (1), Aymaran (1), Barbacoan (2), Boran (2), Cariban (3), Chapacuran (1), Chibchan (2), Chocoan (1), Chonan (2), Guaycuruan (2), Jivaroan (1), Macro-Gêan (4), Nadahup (2), Matacoan (1), Mosetenan (1), Nambikwaran (2), Paezan (1), Panoan (3), Quechuan (3), Tacanan (1), Tucanoan (2), Tupían (9), Yanomaman (1), and unclassified languages or isolates (11).

All in all, the sample represents 25 language families in addition to 11 unclassified languages that possibly constitute their own families. According to the most recent classification of SAILS in Hammarström (2009), there are 111 well-attested language families (or isolates) for SA (whose affiliation is sufficiently demonstrated in publication). This shows that this sample covers only a small part of the linguistic diversity and is therefore to be regarded as such. However, this sample does include members of Hammarström's largest language families: Arawakan with 62 members, Panoan (28), Cariban (32), Quechuan (46), Tupían (76), and Tucanoan (25). It is therefore more precise to say that although this sample does not cover the complete genealogical diversity of SAILS, it does present the whole range from the largest to the smallest families.

Denomination of language families is differentiated from proper languages by adding '-(a)n', such as in Nambikwaran, Nadahup, or Quechuan. This is to prevent ambiguities, for example between Arawak, the language (Lokono Dian), and Arawakan, the name of the language family of which Arawak is a member. An exception is made with Nadahup, which has been coined by Epps (2008, 9-10) to replace the traditionally used but misleading term 'Maku'.

To be consistent, '-(a)n' is used as well where there is no danger of ambiguity, as in e.g. Quechuan: "[t]he use of the denomination "Quechuan", with the ending '-(a)n', has not been common, mainly because there is no particularly variety of Quechuan more entitled to be called "Quechua" than any of the others. The name *Quechuan*, nevertheless, is useful for distinguishing reference to the whole family of Quechuan varieties (languages and dialects) from the use of "Quechua" in reference to individual varieties" (Adelaar 2012, 4, original highlighting).

In addition to the specific language data as shown in table 2.1, several studies of families and regions were consulted. For Tupían this includes e.g. Rodrigues (1999), Rodrigues & Cabral (2012), and Jensen (1998, 1999); for Cariban Gildea (1998, 2012) and Derbyshire (1999); for Arawakan Wise (1986), Derbyshire (1986), Payne (1991), and Aikhenvald (1999); for the Andes Adelaar with Muysken (2004), and Adelaar (2008, 2012), for the Guaporé-Mamoré Crevels & Van der Voort (2008); for the Vaupés Aikhenvald (1999) and Epps (2005); for the Upper Xingú Seki (1999). This is but a short list of available sources and the reader is referred to the references for more.

The questionnaire that was developed for this study (see below), was fully coded for the 63 languages sample, but for certain categories languages were added when data were available. For example, the discussion about Modality in Cariban includes data from Wai Wai and Makushi (section 5.12.4). The list of languages for specific categories is therefore bigger than 63, although the additional languages are not coded in the questionnaire. Hence, the present sample includes 63 SAILs that, besides geographical, genealogical, and typological diversity were also chosen for availability and quality of sources. It could be argued that the language sample is too small to yield significant results, but the number of languages necessary for a valid cross-linguistic study seems to vary from person to person. For example, Aikhenvald (2004, xii), in her study of Evidentiality in a world-wide sample of 500 languages, goes as far as claiming that “five hundred is no more than one-tenth of all human languages [...]. It thus seems most judicious to follow a qualitative approach at the present time, postponing quantitative analysis until more reliable data is available and can be assessed”. I agree that the optimal language sample indeed includes as many languages as possible, but I do believe that an approach with less than the maximally possible number points at least towards tendencies, as for example done in Bybee et al. (1994) with a sample of 76 languages world-wide.

A word about the maps in this study is in order. All maps occurring in the following chapters were created with R (with the exception of map 2.1), an open-source program for graphical and statistical computing.⁷ They are used solely for illustrative purposes. The languages are represented by points whose latitude and longitude values were generated according to Ethnologue (Lewis 2009) and information in the sources. The latitude/longitude values are approximations to those locations. Although in some cases a language is on the brink of extinction by now (e.g. Itonama), it is nevertheless included in the map. For an overview see maps 2.1 and 2.2. Map 2.1 illustrates the position and spread of languages in the sample at the point of contact, based on both linguistic and ethno-historic information (cf. Eriksen 2011, 12-13).⁸ For a better orientation, it additionally gives the political borders of today, as well as geographical denominations used in this study (e.g. Amazonia, Andes, Chaco, etc.). Map 2.2 is computed with R and illustrates the rough position of the languages in the sample. All the maps of the following chapters are based on this model.

⁷ Free download at www.r-project.org. My thanks go to Joshua Birchall, Sander Lestrade and the R mailing list for help with creating maps.

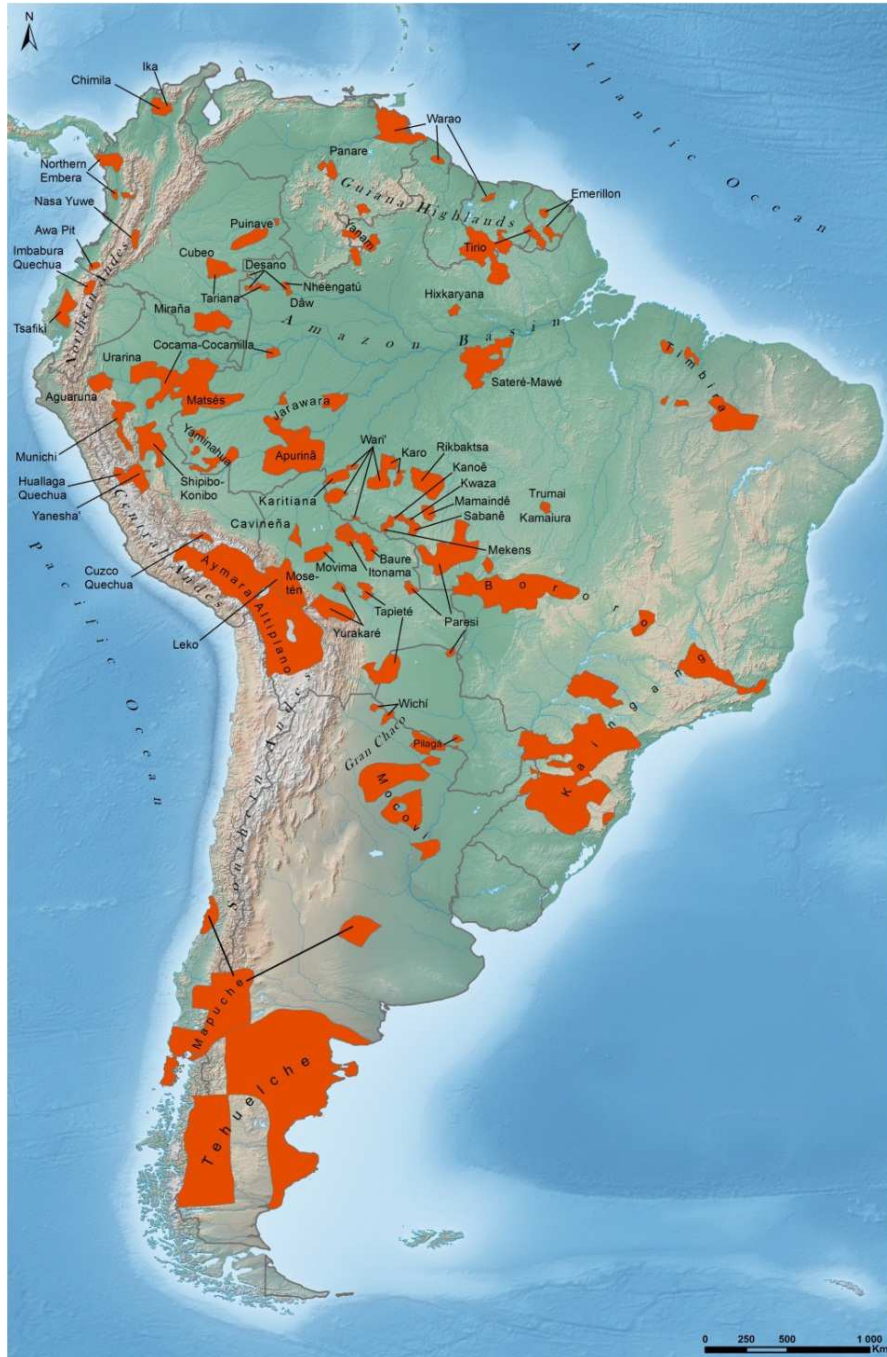
⁸ I am grateful to Love Eriksen who generously provided me with this map.

Table 2.1: List of languages in the sample (languages not included in the original ToC sample in italics)

LANGUAGE	AFFILIATION	PRIMARY SOURCE(S)
Apurinã	Arawakan	Facundes 2000
Baure	Arawakan	Danielsen 2006, Danielsen p.c.
Tariana	Arawakan	Aikhenvald 2003
Yanesha'	Arawakan	Duff-Tripp 1997
Paresi	Arawakan	Brandão 2010, Brandão 2011, Brandão p.c.
Jarawara	Arawan	Dixon 2004
Mapuche	Araucanian	Smeets 2008
Aymara	Aymaran	Hardman 2001, Cerrón-Palomino & Carvajal Carvajal 2009
Awa Pit	Barbacoan	Curnow 1997
Tsafiki	Barbacoan	Dickinson 2002
Miraña	Boran	Seifart 2005, Seifart p.c.
Hixkaryana	Cariban	Derbyshire 1979, 1985
Panare	Cariban	Gildea 1989, 1992, Thomas E. Payne 1990, 1995
Tiriyó	Cariban	Meira 1999, Meira p.c.
Wari'	Chapacuran	Everett & Kern 1997
Chimila	Chibchan	Trillos Amaya 1996
Ika	Chibchan, Aruak	Frank 1985, Frank 1990, Frank p.c.
Embera	Chocoan	Mortensen 1999, Aguirre Licht 1999
Tehuelche	Chonan	Fernández Garay 1998
Mocoví	Guaycuruan	Grondona 1998
Pilagá	Guaycuruan	Vidal 2001
Aguaruna	Jivaroan	Overall 2007, Overall p.c.
Bororo	Macro-Gêan	Crowell 1979, Nonato 2008
Rikbaktsa	Macro-Gêan	Silva 2005, 2011
Kaingang	Macro-Gêan	Cavalcante 1987
Timbira	Macro-Gêan	Alves 2004
Dâw	Nadahup	Martins 2004
Hup	Nadahup	Epps 2005, 2008
Wichí	Matacoan	Terraza 2009
Mosetén	Mosetenan	Sakel 2004, Sakel p.c.
Mamaindê	Nambikwaran	Eberhard 2009
Sabanê	Nambikwaran	Antunes de Araujo 2004
Nasa Yuwe	Paezan	Jung 2008
Matses	Panoan	Fleck 2003
Shipibo-Konibo	Panoan	Valenzuela 2003b
Yaminahua	Panoan	Faust & Loos 2002
Huallaga Quechua	Quechuan	Weber 1996
Imbabura Quechua	Quechuan	Cole 1982, 1985
<i>Cuzco Quechua</i>	Quechuan	Faller 2002, 2003
Cavineña	Tacanan	Guillaume 2008, Guillaume p.c.
Cubeo	Tucanoan	Morse & Maxwell 1999
Desano	Tucanoan	Miller 1999
Karitiana	Tupían	Everett 2006
Karo	Tupían	Gabas Jr. 1999
Mekens	Tupían	Galucio 2001, Galucio p.c.
<i>Sateré-Mawé</i>	Tupían	Silva 2010

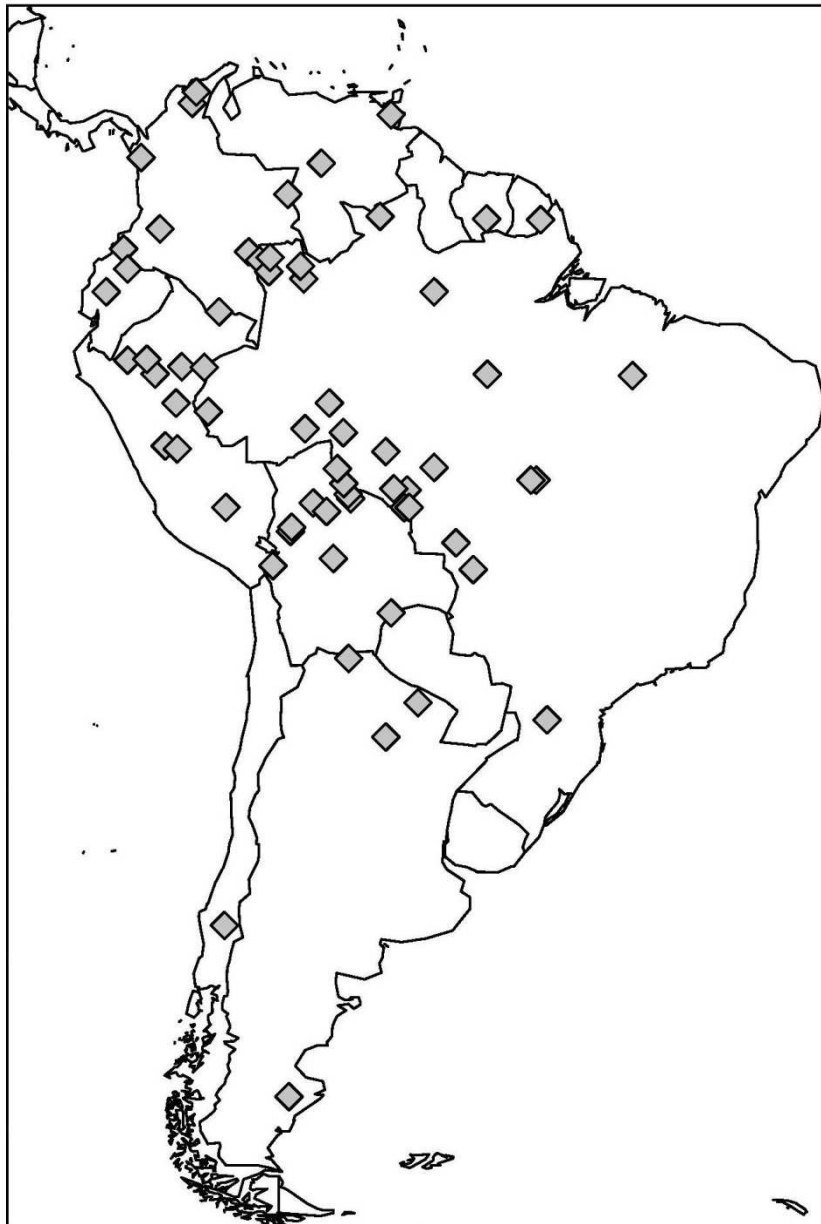
Emérillon	Tupían, Guaraní	Rose 2003
Kamaiurá	Tupían, Guaraní	Seki 2000
Cocama-Cocamilla	Tupían, Guaraní	Vallejos Yopán 2010
Tapiete	Tupían, Guaraní	González 2005
<i>Nheengatú</i>	Tupían, Guaraní	da Cruz 2011
Yanam	Yanomaman	Goodwin Gómez 1990
Itonama	unclassified	Crevels 2006, Crevels p.c.
Leko	unclassified	Van de Kerke 2009, Van de Kerke p.c.
Movima	unclassified	Haude 2006
Yurakaré	unclassified	Van Gijn 2006, Van Gijn p.c.; Gipper 2011
Kanoê	unclassified	Bacelar 2004
Kwaza	unclassified	Van der Voort 2004, Van der Voort p.c.
Trumai	unclassified	Guirardello 1999
Munichi	unclassified	Gibson 1996
Puinave	unclassified	Girón 2008
Urarina	unclassified	Olawsky 2006
Warao	unclassified	Romero-Figueroa 1997

Map 2.1: Approximate position of languages in the sample at point of contact



Map 2.2: Model map of the SAILs sample

Languages in the sample



2.5.2 The questionnaire

The basis for this study is a questionnaire which was developed according to the structure of TAME categories as well as according to the configuration of the complete Traces of Contact database. The final database will be made public in 2013 and all the data will be transparent. Because the individual properties of the features and why they were chosen are discussed in detail in the respective sections, this section concentrates on the structure of the questionnaire. It suffices to say here that, given the temporal limitations of the project, it was attempted to include a mixture of the major categories (e.g. PRESENT for Tense, PERFECTIVE for Aspect, REALIS for Modality, and FIRSTHAND for Evidentiality) and those which may be considered more marginal, but not less interesting (e.g. DESIDERATIVE, FRUSTRATIVE, polite COMMAND).

There are two ways to approach TAME systems: bottom-up or top-down. The bottom-up approach looks at the individual markers and their distribution. It starts with the object of interest and examines its values. This is the ideal case if one attempts a descriptive work, but unfeasible for a questionnaire. The very nature of a questionnaire is establishing clear boundaries between meanings, to demarcate meaning A from B. The very nature of TAME, on the other hand, is that A and B (and C and D and E and F ...) are related to each other and cannot easily be separated. In order to establish some kind of common ground as a basis for comparison, the top-down approach is adopted here: a set of definitions functions as a template and is applied to the data. This has the disadvantage of losing information about markers, but I have commented upon this when necessary in both the study and the questionnaire.

The questionnaire has four topic sections, one each for TAME, amounting to a total of 35 questions. Each question asks for a specific feature and has its own number which codes the section (1 for Tense, 2 for Mood/Modality, 3 for Aspect, and 4 for Evidentiality) and the dependency to other questions, e.g. question 1.2.1 codes for Tense (1), it is dependent of the second question (2), and it is the first to come in line (1). The order of the sections and questions is not important; each independent question can be answered without consulting other questions. A list of all questions is given below:

Tense

- 1.1 Is present tense marked morpho-syntactically?
- 1.2 Is future tense marked morpho-syntactically?
- 1.2.1 In terms of remoteness, how many future tenses are marked morpho-syntactically?
- 1.3 Is past tense marked morpho-syntactically?
- 1.3.1 In terms of remoteness, how many past tenses are marked morpho-syntactically?

Modality

- 2.1 Is realis mood marked morpho-syntactically?
- 2.2 Is irrealis mood marked morpho-syntactically?

- 2.3 Is imperative mood marked morpho-syntactically with one marker for more than one person and number?
- 2.4 Is imperative for 1st person only marked morpho-syntactically?
- 2.5 Is imperative for 2nd person only marked morpho-syntactically?
- 2.6 Is imperative for 3rd person only marked morpho-syntactically?
- 2.7 Is prohibitive marked morpho-syntactically?
- 2.8 Is a polite/ mild/ weak imperative marked morpho-syntactically?
- 2.9 Is intentional marked morpho-syntactically?
- 2.10 Is there a potential modal marker expressing a potential reality (but not necessarily excluding past reference)?
- 2.11 Is certainty marked morpho-syntactically?
- 2.12 Is dubitative marked morpho-syntactically?
- 2.13 Is frustrative marked morpho-syntactically?
- 2.14 Is purposive marked morpho-syntactically?
- 2.15 Is desiderative marked morpho-syntactically?

Aspect

- 3.1 Is perfective marked morpho-syntactically?
- 3.2 Is imperfective marked morpho-syntactically?
- 3.3 Is anterior marked morpho-syntactically?
- 3.4 Is a habitual action marked morpho-syntactically?
- 3.5 Is a continuative action marked morpho-syntactically?
- 3.6 Is an iterative action marked morpho-syntactically? (single and/or multiple repetition)?
- 3.7 Is a completive action marked morpho-syntactically?
- 3.8 Is an incompletive action marked morpho-syntactically?

Evidentiality

- 4.1 Is firsthand information marked morpho-syntactically?
- 4.2 Is non-firsthand information marked morpho-syntactically?
- 4.3 Is secondhand information marked morpho-syntactically (may include third-hand information)?
- 4.4 Is quotative marked morpho-syntactically?
- 4.5 Is visual evidentiality marked morpho-syntactically?
- 4.6 Is inferred evidentiality marked morpho-syntactically?
- 4.7 Is assumed evidentiality marked morpho-syntactically?

The database only shows binary values, i.e. either a language exhibits a morpho-syntactically marked feature, or not. However, for this particular study, multiple values are given. Each question has an answer key which provides possible values. These values differ marginally from question to question, but always have a positive, a negative ('No'), and an

unknown (“?”) value. Positive values include the following options: affix, clitic, particle, auxiliary, stem change/ suppletion, imperative marker and additional marker (question 2.8 only), repetition (of parts) of the verb (Aspect questions only), and numbers in the case of questions 1.2.1 and 1.3.1 (remoteness degrees). A question receives a positive entry when there is morpho-syntactic marking of that feature, negative when there is none, according to source. The questionnaire codes for markers of the value types affixes, clitics, particles, auxiliaries, and stem change in main, non-negative, non-interrogative clauses. The possible values for each answer differ, e.g. ITERATIVE has the additional value of repetition of (parts of) the verb, and the Tense remoteness question values are numbers beginning at zero (i.e. no remoteness degrees). For each question it is also possible to answer with “?” in cases where the sources do not mention that specific feature. Certain Modalities are exceptions to the main clause rule, e.g. IRREALIS markers may appear in subordinate clauses, as do POTENTIAL markers. The PROHIBITIVE of course occurs in negative clauses. For a more detailed discussion look at the respective definition sections. All questions with a positive entry specify the information source (author, date and page number), and negative entries also get a source entry when the absence of that feature is specifically stated.

Due to the final structure of the database, the questionnaire allows for only one answer per question, but there are cases of multiple possible answers. For example, a language that expresses PAST Tense by two markers, e.g. one suffix and one particle, cannot have both answers in the value column. The marker that is most grammaticalized is entered, with the preference going from affix (most grammaticalized) over clitic to particle and auxiliary. I have kept track of alternative ways of marking in the questionnaire in the comment section. One feature marked by two morphemes should be rare, however. Often one marker has a slightly different meaning, so for the questionnaire one should carefully consider the environments both markers occur in. What seems like two PAST markers may be one PAST marker and one PERFECTIVE marker. Or one of both markers is the general one and the other only added to convey remoteness degrees. Or they are in complementary distribution, one in assertive clauses, and one in interrogative clauses. The choice made is in favor of degree of grammaticalization, dominance (in terms of frequency or productivity), and occurrence in assertive, positive main clauses (some Modal categories being exceptions).

The following four chapters constitute the main body of the study. They first present the definitions of the TAME categories used in this study and then present the results, with a focus on geographical and genealogical distributions. They are arranged as follows: Tense (chapter 3), Aspect (chapter 4), Modality (chapter 5), and Evidentiality (chapter 6). They are ordered to reflect the order in the questionnaire and can be read separately without having to consult the others, with the exception of the stability sections. Section 3.10 introduces the concept of temporal stability necessary to understand the stability sections in the other chapters. In order to fully understand the points of overlap I recommend reading all chapters in the order they are given.

3. TENSE

3.1 INTRODUCTION

Klein (2009, 40–41) claims that languages can express time in six different ways: through Tense, Aspect, Aktionsart, temporal adverbials, temporal particles, and discourse principles. Of these, the focus of this chapter is on Tense: “[t]ense is a grammatical category of the verb in its traditional understanding, it serves to locate the situation in relation to the “now” of the speech act” (ibid. 40). Aspect will be dealt with in the following chapter, but because Tense and Aspect are highly related, it will be mentioned where relevant here. Aktionsart, temporal adverbs, and discourse principles are not coded in the questionnaire and therefore not or only partly discussed here, but indeed there would be enough material to allow for several dissertations on each of these. Since particles do fall under my definition of morpho-syntactic marking (see chapter 2), they are included in this study.

If a language has morpho-syntactic Tense marking (obligatory or not) it is Tensed, and if it does not it is Tenseless (cf. Smith 2005, Stassen 1997, Comrie 1985 etc. for similar classifications). This does not entail that a language has no means of expressing time but simply that it lacks such morpho-syntactic marking. Tensed languages in this sample can be differentiated according to which distinctions they make:

1. Two-way distinctions (such as FUTURE/ NONFUTURE and PAST/ NONPAST),
2. Three-way distinctions with the full set of three overtly marked absolute Tenses,
3. FUTURE/ PAST marking.

The latter is not a type of two-way distinctions; FUTURE/ NONFUTURE and PAST/ NONPAST include the possibility that PRESENT is marked (in a cumulative morpheme with either FUTURE or PAST), whereas as FUTURE/ PAST marking languages never overtly mark PRESENT. If they would, there would be a three-way distinction. The section about Tenseless languages (3.3) includes a discussion of those languages that do not show any overt marking characteristic of three groups above. Tensed languages can also be grouped according to how many remoteness distinctions they make; this is discussed in section 3.6.

The following sections discuss these systems in detail, but before a word about the use of the term “system” is necessary. A system denotes a general semantic concept and does not refer to a specific syntactic structure, such as order of affixes. If a language has one verbal slot for Tense and another for Modality, and the FUTURE marker is in the Modal slot, it is nevertheless treated as Tense as long as this is the primary meaning of the element (for the problematic relationship of FUTURE with Modality see below section 3.7).

This chapter starts with the working definitions of Tense and remoteness degrees (sections 3.2.1 and 3.2.2). Detailed analyses of Tenseless languages (3.3), the different Tense systems occurring in the sample (3.4), and the languages which mark the maximal number of Tenses (3.5) follow. Section 3.6 focuses on remoteness degrees. Sections 3.7 to 3.9 present first the typological distribution in the sample (3.7) and then focus on the geographical (3.8)

and genealogical (3.9) distribution of Tense marking. The typological stability of Tense is discussed in section 3.10, and a summary is given in 3.11.

3.2 DEFINITIONS

3.2.1 PRESENT, PAST, and FUTURE Tense

The approach to Tense outlined here relies on the groundbreaking study of Hans Reichenbach in the beginning of the 20th century. Reichenbach approached the notion of Tense from a philosophical, empirical, and logical point of view and presented an accessible as well as logical theory that is consulted by typological linguists and scholars from other disciplines until today (see e.g. Annerieke Boland (2006, 52ff.) for an introduction to Tense). The approaches to time and Tense are manifold and not restricted to the field of linguistics. In the present survey, I will follow a linguistic approach but it should be borne in mind that time in linguistics is hardly separable from time in other disciplines whereas Tense is and that Tense is a very specialized and abstract concept which helps us dealing with the grammatical expression of time in languages but does not represent an ultimate or complete solution. The following sections introduce the basic principles needed to linguistically define Tense; it first outlines PRESENT, PAST, and FUTURE and then illustrates remoteness degrees of PAST and FUTURE (3.2.2).

Reichenbach (1948) discusses the existence of the denominations “point of speech” (S), i.e. the moment an utterance is spoken, and the “point of event” (E), i.e. the moment the action of the utterance takes place. Together they show a relationship in time: the event took place before the point of speech (PAST), takes place simultaneously (PRESENT), or takes place after it (FUTURE), in Reichenbach’s terms being anterior, simple, or posterior (Reichenbach 1948, 297). Consider the following examples taken from Kwaza:

(3.1) Kwaza (unclassified; Van der Voort 2004, 102, 390, 400)

- (a) *hã 'kui-da-ki*
water drink-1SG-DECL
'I am drinking/drunk water.'
- (b) *kukui'hỹ-da-da-ky-hỹ-ki*
ill-1SG-1SG-PST-NMLZ-DECL
'I was ill.'
- (c) *ja-'nã-da-ki*
eat-FUT-1SG-DECL
'I'm going to eat.'

Kwaza does not have a morpho-syntactic PRESENT marker and the PAST marker is not obligatory, therefore (3.1a) has both PRESENT and PAST as possible reading (ignore the PROGRESSIVE here). The first example shows a typical PRESENT Tense, the second a PAST Tense and the third a FUTURE Tense. In all these examples the points of event relate in a special way to the points of speech. A Tense always consists of the relationships between these

points, including the possibility that they overlap. However, this explanation fails to provide for Tenses ancillary to the past-present-future continuum:

- (3.2) Imbabura Quechua (Quechuan; Cole 1982, 143)
Marya Agatu-pi kawsa-shka-ta kri-rka-ni
 María Agato-in live-PST.NOM-ACC believe-PST-1
 'I believed that María had lived in Agato.'

In (3.2), there are two different events in relation to the speech time, the believing of the speaker and the living of Mary in Agato, and, most importantly, they have different PAST Tenses in relation to the point of speech. The example expresses that the speaker believed at a certain point in time in the past that a person called Mary had lived in Agato at a time that is even further removed from the speech moment. So firstly, there seems to exist more than one PAST, and secondly, the fact that the Tense of the subordinate clause seems to be in some kind of relationship with the Tense of the main verb rather than with the speech time suggests a shortcoming of the two-dimensional approach. Reichenbach proposes that a third point in time is needed to fully account for the whole range of Tenses: the point of reference (R). Applied to (3.2) the point of reference is the time of believing which took place before the speech time, a simple PAST, and the point of event took place even before the point of reference, thus generating a PAST in the PAST. Any Tense marker always refers to the relationships between these three points, including the possibility that two of those points overlap. In fact, the congruence of two points often makes them look like one, leading to the outdated belief that the Tense continuum is two-dimensional. The following examples are taken from Reichenbach to illustrate this: In the sentence 'I saw John' the points of reference and event are the same, and they are before the point of speech, generating simple PAST. In 'I had seen John' the point of event is still before the point of speech, therefore PAST, but in between them now there is another event, the point of reference. The speaker is probably talking about an event, i.e. the point of reference, which happened before the moment of speech, and in relation to that event he did see John, generating a PAST in the PAST. Thus the point of reference differs from the point of event. In 'I will have seen John' the speech moment is before the point of reference, and the event is in relation to R located in the past, from the point of S is still in the future, generating a PAST in the FUTURE. In fact, all three points may overlap, and then they indicate the PRESENT, as in 'I see John'. The point of speech, the reference point, and the point of event overlap. Therefore we logically deal with a three-dimensional Tense system, where the relationships of point of speech to point of event, of point of event to point of reference, and of point of reference to point of speech all have to be taken into account (see figure 3.1).

Comrie (1985, 9) defines Tense as "grammaticalized expression of location in time". In his opinion, similarly to Reichenbach, Tense relates the time of the situation with the time of the utterance, the term *situation* referring to "processes, events, states etc." (ibid. 5). Time can be represented as a straight line, from left to right, with an arbitrary point representing

the present, and past and future being left or right respectively. Tense now is the relation between any two points along this time line. (There are obvious flaws with the representation of time in a straight line, as discussed in Comrie (*ibid.*), but it will be used here nevertheless for the sake of visual demonstration.) Comrie uses a reference point which he calls deictic center. His argument is that Tense is a deictic category, i.e. Tense needs a central point around which the temporal relations revolve. When the deictic center falls into the present moment and this is not specifically expressed, it generates an absolute Tense. When the reference point is not specifically limited to the present time, i.e. may have present as one possibility (depending on context, interpretation, etc.), it generates relative Tense. It is important to realize that the reference point has to be specifically stated in relative Tense and may be on any location on the time line, including the present moment, whereas absolute Tense has an overlap of present moment and reference point without specifically marking it. Therefore the absolute/relative Tense distinction is not one of having the reference point coinciding with the present or not, but rather whether this is explicitly stated (*ibid.* 58). The PAST/PRESENT/FUTURE distinction therefore is basic absolute Tense as the reference point is inherently present. Comrie differs from Reichenbach in that for a simple PAST Tense Reichenbach puts the point of reference at the point of event, rather than at the point of speech. For Comrie, an absolute Tense always has the point of reference and point of speech coinciding.

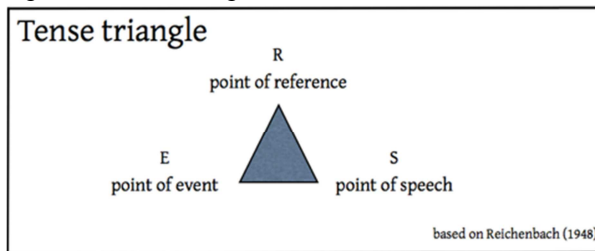
Other researchers also make a distinction between absolute and relative Tense, but define it slightly differently. For example, Chung & Timberlake (1985, 203) take an overlap of speech moment and Tense locus as absolute, and any other Tense where the Tense locus does not overlap with the speech moment is relative: “Tense systems (or subsystems) in which the speech moment serves as the tense locus are traditionally called absolute tense; systems (or subsystems) in which some other point is the tense locus are called relative tense”. Bhat (1999, 14) does the same, labeling absolute deictic and relative non-deictic Tense, and deictic Tenses having the utterance time as reference point and non-deictic having some other time as reference point. Bhat seemingly disagrees here with Comrie who argues that Tense is always deictic, i.e. “relating entities to a reference point” (*ibid.*), including relative Tenses. Bhat, however, defines deictic as denoting “an event which is connected with the speech act” (*ibid.*), so any event that has a different reference point than the speech moment is non-deictic in his terminology.

Givón (2001, 286) also employs the three-points-concept, but introduces another term, the “anchoring” of time. The event is anchored, i.e. fixed, on a certain point in time by the reference point, and when that overlaps with the speech moment, it is absolute Tense. On the other hand, relative Tense is when the reference point is anchored in the past or future to the time of speech. As we can see, Comrie, Givón, and Bhat agree with Reichenbach that three points in time are necessary for a comprehensive theory of Tense, and this concept is widely agreed upon by other scholars as well. The Kwaza examples (3.1a) - (3.1.c) above are absolute Tenses, and (3.2) is relative Tense. The absolute Tenses are fixed on the time line. The relative Tenses are not fixed as their anchor time can be different for each utterance.

Both can have the PRESENT as the anchor time, but as Comrie (1985, 58) points out: “the difference between absolute and relative tense is not that between the present moment versus some other point in time as reference point, but rather between a form whose meaning specifies the present moment as reference point and a form whose meaning does not specify that the present moment must be its reference point. Relative tenses thus have the present moment as one of their possible reference points”. Because the present study only deals with absolute Tense it is not necessary to further dwell on relative Tense here.

The different terminology employed by different scholars adds to a general confusion but expresses the same core elements. The only major difference is the one of whether relative tenses may or must not have the present as reference point, and that is not important here as the questionnaire only covers absolute Tense. I agree with Comrie that Tense is a deictic category, as any Tense, whether absolute or relative, is a representation of the relationship between three points in time that revolve around a center. This can be shown in the following figure:

Figure 3.1: Tense triangle



Because this study is concerned with absolute Tenses, and because absolute Tenses always have R and S coinciding, we can ignore the point of reference. Subsequently, any relationship between the points of event and speech characterizes an absolute Tense. The following figure gives an overview of these relationships:

Figure 3.2: Absolute Tense illustrated (based on Reichenbach 1948 and Comrie 1985)

Absolute Past	E<S	I saw John.
Absolute Present	E=S	I see John.
Absolute Future	S<E	I will see John.
(< precedes, = coincides with)		

The Tense section of the questionnaire is designed to give information about absolute Tense marking, but not relative. Sometimes, when one marker can mark both absolute and relative Tense, then relative Tense is automatically coded as well. Such a marker is for

example the suffix *-shita* which marks FUTURE in Yurakaré in absolute as well as relative Tense:

- (3.3) Yurakaré (unclassified; Van Gijn 2006, 181, 182)
- (a) *tishilë mi-la-bobo-shita-tu* FUTURE
 now 2SG-AFF.OBJ-hit/kill-FUT-1PL.SBJ
 ‘Now we are going to kill him on you!’
- (b) *a-ushpë-Ø=w=ti ma-che-shita-Ø* FUTURE in the PAST
 INCL-bathe-3=PL=DSBJ 3PL-eat-FUT-3
 ‘When they were bathing, he was going to eat them.’

The three-points-in-time definition is the most basic approach to Tense as possible and it deliberately does not address the problems that arise from the grey zone between Tense and Aspect, Modality, and Evidentiality. Nevertheless, these problems arise when coding the languages. None of the Tenses is without catches on to which Aspectual, Modal, or Evidential meanings may hook, making it difficult to decide which category the marker really belongs to. In those cases, one must go back to the dominance principle mentioned in chapter 2. For example: A marker may look to be a prototypical FUTURE, denoting events that take place after the moment of utterance. But at a second glance, one discovers that it only applies to those events in the future that are planned or intended. That makes it a modal marker of INTENTION with FUTURE as a secondary (inherent) meaning. Another marker may look like a nice and neat PAST marker only to reveal that it is actually coding events that happened in the past but still have present relevance, e.g. being an ANTERIOR marker. In that case it is crucial to remember that markers are Tense markers when their predominant function is coding the relationship of three points in time on the time line.

As has been pointed out in chapter 2, obligatoriness is not a necessary characteristic of a TAME marker in this study. This in contrast to the opinion of several authors who claim that a grammaticalized expression of time is automatically obligatory, or it would not be grammaticalized. Advocates of this opinion are e.g. Comrie (1985), Dahl (1985), Smith (2008) and Stassen (1997). While I agree that grammaticalized markers show a certain tendency to be obligatory, the notion of obligatoriness is too unspecific to be of value here. For example, the border between obligatory and non-obligatory is a matter of dispute: does an obligatory marker have to occur in every possible instance? Is a marker still obligatory when it occurs in main, but not in subordinate clauses? Is it still obligatory when it occurs only once in the beginning of a paragraph although it clearly has scope over the whole paragraph? These questions are not yet answered satisfactorily. Additionally, it is impossible to carry out obligatoriness studies on all potential TAME markers in the sample, and the sources themselves rarely mention obligatoriness at all.

Additionally to the question whether or not a Tense marker has to be obligatory another problem arises: that of the position of a morpheme in the clause. Especially with regard to Tense this is a sensitive topic as the ongoing dispute about nominal Tense shows

(c.f. Nordlinger & Sadler 2004 and 2008, Tonhauser 2007, Muysken 2008). Because the focus of this thesis is on semantics the position of the morpheme in question does not matter as long as the temporal meaning refers to the verbal action specified in the same clause. In cases where one and the same marker has scope over a nominal constituent or the action of the clause it is still treated as Tense marker, but when a marker occurs exclusively with scope over a nominal constituent it is not. For example, the PAST enclitic =*ene* in Paresi can attach to both verbs and nominals with scope over the respective constituent; because it can have scope over the verb it is a Tense marker. But in Mamaindê there is a set of suffixes that is restricted to occur on nominals with no scope over the verbal action (Eberhard 2009, 343) and this set is not considered Tense here.

(3.4) Paresi (Arawakan; Brandão 2011)

- (a) *zane maidyat-ita=ene*
 go fish-PROG=PST
 'He was fishing.'
- (b) *mama=ene atyo ka-eare-tya natyo no-zoemi-nita*
 mother=PST TOP ATTR-name-VBLZ 1SG 1SG-child-?
 'It was my late mother who gave me a nickname.'

(3.5) Mamaindê (Nambikwaran, Eberhard 2009, 343, 371)

- (a) *ta-walek^han-leʔi-tu*
 POSS1-chief-NOM.PST-FNS
 'my past chief'
- (b) *nūsa-jahon-sihatija-tu kajat-kanīn*
 POSS1.PL-old.man-NOM.PST-FNS CORN-NCL.ROUND
haʔfin-jeʔ-Ø-nīnta-wa
 plant-EMP-SBJ3-GKN-DECL
 'Our ancestral men planted corn seeds (as everyone knows).'

The following tables present an overview of the terminology of Tenses used in this in comparison to other studies (tables 3.1 and 3.2):

Table 3.1: Terminology for time continuum:

Author	Terminology		
Reichenbach 1948	Point of speech	Point of reference	Point of event
Comrie 1985	Speech situation	Reference point	Situation
Timberlake & Chung 1985	Speech moment	Tense locus	Event frame
Bhat 1999	Utterance time	Reference point	Event or situation
Givón 2001	Time of speech	Reference time	Event time
Ultan 1978	Moment of speech	Relative time	Moment of speech
Mueller (present volume)	Point of speech	Point of reference	Point of event

Table 3.2: Terminology for absolute Tenses:

Author	PAST	PRESENT	FUTURE
Reichenbach 1948	Before	Simultaneous	After
Comrie 1985	Before	Simultaneous	After
Timberlake & Chung 1985	Anterior, prior	Simultaneous, overlapping	Posterior, subsequent
Bhat 1999	Before	Simultaneous	After
Givón 2001	Precedes	Right at	Follows
Ultan 1978	Antecedent	Coincide	Subsequent
Mueller (present volume)	Before	Simultaneous	After

3.2.2 Remoteness degrees

Additional to just coding an event that is removed from the point of speech (FUTURE or PAST) languages may code for the time length elapsing between the point of event and the point of speech, i.e. the remoteness degree. Naturally, this applies only to PAST and FUTURE. These distinctions represent events that are removed from the point of speech, varying from recent or immediate over remote to historical or even mythical (in the case of PAST). For example, Karo has two (non-obligatory) particles expressing PAST: *co* for recent or general PAST and *kán* for remote/ mythical PAST:

(3.6) Karo (Tupían; Gabas Jr. 1999, 175)

- (a) *púŋ* *o=ʔe-t* *co*
 shoot 1SG=AUX=IND PST
 'I shot.'
- (b) *õn* *ĩriŋ* *ma-céri-t* *kán*
 1SG girl CAUS-heal-IND REM.PST
 'I healed the girl (long ago).'

The recent PAST in Karo can refer to time length between a minute ago to a year ago, remote to anything from ten to hundred years ago. The cutoff point between when recent stops and remote begins is often blurred and depending on the judgment of the speaker. That is why the questionnaire asks for how many distinctions are made instead of which exactly. It will be shown below (section 3.6.9) (i) that remoteness degrees are subjective and that (ii) recent, remote, and similar terms are used for a multitude of different time spans in different languages and therefore not *per se* comparable.

To summarize: Absolute Tense as used in this study is the grammaticalized expression of time, but not necessarily obligatory. It is a deictic category showing the relationship of two points on a time line (point of event, point of speech). FUTURE, PAST, and PRESENT can be defined by the combination of these points: FUTURE is an event that is located after the point of speech, PAST is an event that is located before the point of speech, and PRESENT is an event simultaneous with the point of speech. It was argued that because the notion of a point of reference is not of consequence in absolute Tense it is not needed in this study. Events

removed from the PRESENT can cover different time spans, from near over remote to far (either in the PAST OR FUTURE).

3.3 TENSELESS LANGUAGES

A number of languages in the SAILS questionnaire have a negative entry for the Tense questions; they seem to be Tenseless, i.e. do not have morpho-syntactic marking of absolute PRESENT, FUTURE OR PAST, or remoteness degrees in the latter two. These are Baure, Tsafiki, Pilagá, Mocoví, Nasa Yuwe, and Yanam. Additionally, there are three borderline cases: Urarina, Mapuche, and Trumai, which are treated as being Tenseless, although they seemingly have some morpho-syntactic Tense marking. However, they are not comparable to the Tense marking in the other SAILS and thus discussed in this section as well.

Comrie (1985, 52) already noticed that Burmese and Dyrbal lack grammatical Tense marking, rather having “a basically modal opposition [that] has implications for the time reference without this time reference being grammaticalized in the language”. Comrie also points out that languages which do have morpho-syntactic Tense marking allow for constructions without Tense (*ibid.*). Both of these phenomena, here called complete and partial Tenselessness, occur in this sample and are discussed in this section together with the implications for a general theory of Tense. Smith (2005) proposed an essentially three-way Tense classification: fully Tensed, mixed-temporal, and Tenseless languages. The first two categories differ in whether or not they mark Tense obligatorily; as this distinction is not made in the present sample, all languages that are not discussed in this section must belong to either one of them. Following is a typological overview of selected Tenseless SAILS and borderline cases, and how they do express temporal relations instead.

Pilagá has no morpho-syntactic Tense marking, but a number of lexical expressions for indicating time, e.g. *sekaet* ‘yesterday’ (Vidal 2001, 228). Beside time lexemes, Pilagá features deictic classifiers and a large number of directionality suffixes, some of them with deictic reference (*ibid.* 235). Because Tense is a deictic category, it could theoretically be the case that spatial reference also expresses temporal reference, and indeed this seems to be the case in Pilagá. The spatial classifier *so* ‘going away’ can not only refer to the movement of the classified entity, but also to the fact that the action of the same clause is in the past (*ibid.* 119–120). As for the deictic directional suffixes, Vidal presents examples with e.g. directional *-ge* ‘thither’ and past time reference, but it remains unclear whether the temporal interpretation relies on the directional or on context.

(3.7) Pilagá (Guaycuruan; Vidal 2001, 336, 120, 237)

- (a) *weta-ge' sekaet da' Formosa ñi Xuan*
 be-DIR.thither yesterday CLF Formosa.city CLF John
 ‘John was in Formosa city yesterday.’
- (b) *an-sa-nem so' paan*
 2OBJ-1-give CLF.going.away bread
 ‘I gave you bread.’ (The bread is not in view any more.)

- (c) *na-č̣a-ge'*
 3-go-DIR.thither
 'He went away.'

Similarly to Pilagá, Yanam utilizes a number of features for temporal reference: "Temporal reference [...] is expressed by a combination of aspectual morphemes, temporal adverbs, and the context. [...] The time when an event occurs or occurred becomes clear only after temporal adverbs and the context are fully taken into consideration" (Gómez 1990, 93).

- (3.8) Yanam (Yanomaman; Gómez 1990, 84)
hena th@h@ ora ham camak hl-I
 early when forward DIR we go-IPVF
 'Tomorrow we are going upstream.'

Baure employs a range of temporal adverbs (Danielsen 2007, 303ff.) as well as Aspect marking for indicating time. For example, the PERFECT preverbal particle *ver* occurs in clauses with both present and past reference, depending on the verb (ibid. 274). The INTENTION suffix *-pa* is often used for future reference, either absolute or marking a future in the past (ibid. 264), as well as the INTENTION particle *kač* (ibid. 277), and the terminative particle *eto* seems to be expressing exclusively past time reference. According to Krasnoukhova (2012, 241), Baure is one of the languages in SA which encode time in noun phrases.

Tense marking in the three borderline cases Urarina, Trumai, and Mapuche do not fit the classification set up in 3.2, but they are not completely Tenseless either. In Urarina, future time reference can be expressed by the combination of the IRREALIS marker and assertive. But Urarina has another way to express time, i.e. by the general remoteness enclitic *=lu* that usually applies to remote PAST but can also refer to remote FUTURE depending on context and further temporal specification in the clause. The remoteness enclitic is apparently not obligatory for an action in the past that is remote, as shown below.

- (3.9) Urarina (unclassified; Olawsky 2006, 482, 435)
- (a) *kaa dzaura+atane itahe-ri-teãu=ni=tea*
 this flesh+land destroy-IRR-1SG=ASSR=EMP
 'I will destroy this earth.'
- (b) *nitoaneĩ hetau=te kaṭea leṃu-e=lu lomaj,*
 like.that HEA=FOC man sink-3=lu Lomai
edara ne-ĩ ku-wru-a=ne kujna
 water.people be-PTCP go-PL-3=SUB so.that
 'Lomai sank the people like that, so that they would become water people.'

The remoteness enclitic can form a temporal adverb together with the reassurance enclitic =*tau*: *taulu* ‘earlier’ (Olawsky 2006, 505). So although Urarina has one marker that does refer to temporal remoteness, it is treated here as being Tenseless, because it is not a general Tense marker.

Trumai is a similar borderline case. It marks Tense by special Tense/ focus particles. The particles *ka in* and *chi in* are Tense/ focus particles, expressing new information, emphasis etc., but *ka in* also marks for PRESENT/ recent PAST, and *chi in* marks for distant PAST (Guirardello 1999, 170ff.). Thus, Trumai marks PRESENT and PAST, with an overlap between recent PAST and PRESENT, but not FUTURE. This system almost looks like a FUTURE/ NONFUTURE system, with unmarked FUTURE. However, Trumai NONFUTURE is ambiguous only between PRESENT and recent PAST, whereas distant PAST is clearly distinguished.

Mapuche at first glance seems to have a PAST/ NONPAST Tense system, but a closer look reveals that there is no clear opposition in Tense. Mapuche morpho-syntactically marks PAST Tense and pluperfect, but at least the former is not obligatory. The PAST particle *nga* occurs in clauses with past time reference, but also expresses sadness or regret (Smeets 2008, 331). The pluperfect, which is not coded in the questionnaire, is marked by two suffixes that are in complementary distribution: *-mu* and *-(ü)wye*. Both mark a point in time before another point in time (PAST) and indicate that the state talked about is no longer true (ibid. 331, 254). In general, Tense marking plays a minor role in Mapuche, which tends to express temporal relationships by Aspectual and Modal suffixes/ particles and their interaction. So even when there is no PAST marker, there seems to be no necessity to interpret the clause as PRESENT or FUTURE, but it can nevertheless be PAST. Temporal meaning rather relies on the presence of certain Aspectual and Modal markers, e.g. the unmarked PERFECTIVE can indicate PAST, as in (3.10a).

(3.10) Mapuche (Araucanian; Smeets 2008, 165, 230, 331)

- (a) *lef-üy*
run-IND
‘he ran’
- (b) *epu antü-nge-y ñi kutran-küle-mu-m*
two day-VBLZ-IND-3 POSS.1SG illness-STA-PLPF-IVN
‘Two days ago I was ill (my having been sick is two days ago).’
- (c) *may, chümül nge-me-n nga*
yes, once be-THI-IND.1.SG PTC
‘Yes, I have been there once.’

While some languages are completely Tenseless, many SAILS actually are partially Tenseless. They may mark Tense (obligatorily or not), but still allow for constructions that do not have Tense marking. Comrie (1985, 52–53) points out that for example in English no Tense marking is shown on derived nominals, on infinitives after a finite verb, or on apparently redundant verbs. Similarly Wari’, although having a basic FUTURE/ NONFUTURE

binary split (see section 3.4.2), allows for clauses that are marked with a separate set of Tenseless markers (Everett & Kern 1997, 310). Some morphemes cannot be used with Tense markers but require Tenseless verbal inflectional clitics, such as the negative *a'* / *'ara*.

(3.11) Wari' (Chapacuran; Everett & Kern 1997, 38)

<i>querec</i>	<i>'a</i>	<i>tocwa</i>	<i>wari'</i>
see	NEG.SG	PASS.3SG.M	person

'The person was not seen.'

Tenselessness is not only restricted to certain constructions, but Tense marking can also be dropped in constructions that otherwise are marked. In Jarawara, Tense marking is not obligatory, and often a Tense marked clause is repeated without overt Tense marking (Dixon 2004, 196).

(3.12) Jarawara (Arawan; Dixon 2004, 195)

<i>faa</i>	<i>mee</i>	<i>wisa</i>	<i>na-ra-ke</i>	<i>fahi,</i>
water.F	3N_SG.SBJ	bail.out	AUX-IMM.PST.EYE.F-DECL.F	THERE.N_VSB
<i>faa</i>	<i>mee</i>	<i>wisa</i>	<i>na</i>	
water.F	3N_SG.SBJ	bail.out	AUX.F	

'They bailed out water (from the canoe) there, they bail out water.'

Cavineña even allows for bare stems without Tense inflection, although Tense marking is obligatory (Guillaume 2008, 179). These are just a few examples, and a study of the phenomenon of Tense dropping, perhaps also in relation to Tenselessness, would certainly be very fruitful.

We have seen that Tenseless languages are lacking overt morpho-syntactic Tense markers, but are not devoid of expressing temporal reference. This does not only include temporal lexemes and deictic spatial markers, but another category that is structurally close to Tense also takes a hand in expressing temporal information: Aspect. Smith (2005, 30) observes that in languages without obligatory Tense marking "temporal location is inferred from aspectual information". This is true for this sample as well.

3.4 TENSE SYSTEMS

3.4.1 Overview

Comrie (1985, 49) already noticed that many European languages have split Tense systems: "many languages in fact have a basic two-way split, with either an opposition between past and non-past or between future and non-future". The same can be said for the SAILs in this sample, with FUTURE/ NONFUTURE systems being most frequent. Interestingly, the also theoretically possible split between PRESENT/ NONPRESENT does not occur. Comrie uses the term 'binary split' for any kind of binary opposition, but the following section differentiates between a simple and a binary split, explained below. Split systems can have various

incarnations: marking for FUTURE, i.e. a clause unmarked with that marker is automatically PRESENT OR PAST; this may entail that FUTURE marking is obligatory. Or marking for FUTURE and for PAST/ PRESENT with two different markers. The same holds for PAST/ NONPAST: either only PAST is overtly marked, with PRESENT and FUTURE being unmarked, or both PAST and FUTURE/ PRESENT are marked. This can be further divided into systems that are binary; i.e. distinctions are overtly marked for both parts of the split, or systems that mark only one part and the other one is simply unmarked and ambiguous. Any apparently unmarked category qualifies as zero marker if it fulfills the criteria outlined by Bybee (1994), the most important being that marking is obligatory, i.e. when the absence of a marker directly points towards the absence of a specific meaning and the presence of another. Unfortunately, there is often little or no information to be found in the sources, or the terms ‘unmarked’ and ‘zero marking’ are used as synonyms, making it difficult to establish the status of an unmarked category. Additionally, languages are often not strict about obligatoriness, a factor crucial to establishing zero markers. They may have an obligatory FUTURE marker indicating a zero PAST/ PRESENT morpheme, but in the actual usage it is revealed that a sentence unmarked for Tense is not obligatorily PAST/PRESENT marked, but may well have FUTURE reference. For the present study this means that languages that apparently have a simple split Tense system may well turn out to have a binary split, but that requires further research and is not done in this study.

Two major patterns catch the eye: Firstly, no language marks a simple split of unmarked FUTURE/ marked NONFUTURE or unmarked PAST/ marked NONPAST, or a simple split with unmarked NONPAST/ marked PAST. All PAST Tense systems occur only as a binary split, i.e. there are two overt markers for PAST and NONPAST. But secondly, languages with FUTURE simple split systems do occur and outweigh the FUTURE binary split by 14:3 (see table 3.3).

Ultan (1978, 88) grouped the languages in his sample into prospective and retrospective; the former being similar to a PAST/NONPAST distinction and the latter to a FUTURE/ NONFUTURE distinction: „Prospective: If a present tense may ordinarily mark [...] future of the latter may be unmarked [...]. Retrospective: If a present tense may ordinarily mark [...] past or if the latter may be unmarked”. There are slightly more prospective than retrospective languages in his sample, but interestingly most of the retrospective languages are American Indian languages (he does not say which) whereas most Indo-European languages are prospective (ibid.). This coincides with the present results where FUTURE/ NONFUTURE Tense systems (i.e. similar to retrospective) occur most frequently within two way splits.

The following paragraphs present the occurring Tense systems in the sample and give illustrative examples.

Table 3.3: Tense systems

Tenseless	Baure, Tsafiki, Pilagá, Mocoví, Nasa Yuwe, Yanam (Urarina, Mapuche, Trumai)			
Tensed	Two-way split	simple split	FUTURE/ NONFUTURE (marked/ unmarked)	Apurinã, Yanesha', Chimila, Bororo, Kaingang, Timbira, Kamaiurá, Itonama, Yurakaré, Kanoê, Munichí, Emérillon, Nheengatú, Sateré-Mawé
		binary split	PAST/ NONPAST (marked/ marked)	Warao, Hixkaryana, Matses, Panare, Movima, Rikbaktsa
			FUTURE/ NONFUTURE (marked/ marked)	Aymara, Karitiana, Wari'
	three-way split (marked/ marked/ marked)			Tariana, Tiriyó, Sabanê, Desano, Embera, Leko, Aguaruna, Mamaindê, Cubeo
	FUTURE/ PAST (marked/ marked)			Paresí, Jarawara, Awa Pit, Ika, Mirañá, Tehuelche, Dâw, Hup, Wichí, Mosetén, Shipibo-Konibo, Yaminahua, Huallaga Quechua, Cuzco Quechua, Imbabura Quechua, Cavineña, Karo, Mekens, Cocama-Cocamilla, Tapiete, Kwaza, Puinave

3.4.2 Binary split FUTURE/ NONFUTURE

Languages overtly marked for both FUTURE and NONFUTURE (binary split) are Aymara, Karitiana, and Wari'. All have overt ambiguous PRESENT/PAST (NONFUTURE) markers and distinct FUTURE marking. Karitiana makes a distinction of Tense marking within affirmative clause types: for affirmative clauses of the types copular and speech act participant verbs stems ending in a vowel receive suffix *-t* for NONFUTURE events. If there is an Aspectual marker in the clause as well, the event is usually PRESENT, if not, usually PAST. Vowel-final verbs receive suffix *-j* for FUTURE events, and *-i* for consonant-final verbs. There is a different set of suffixes for Tense marking in "Verb-focus" clauses; consonant-final verbs receive suffix *-in*, vowel-final verbs suffix *-n* for NONFUTURE, and suffix *-ndaki* for FUTURE. Thus, although Karitiana has different sets of markers according to clause type, the general binary FUTURE/ NONFUTURE split is maintained (Everett 2006, 265ff.).

(3.13) Karitiana (Tupían; Everett 2006, 268, 269)

- (a) *i-ta-seʔi-t*
 1SG.ABS-SAP-drink-NFUT
 ‘I drank.’ or ‘I’m drinking.’
- (b) *i-taka-karĩna-j*
 1SG.ABS-SAP-turn-FUT
 ‘I will turn.’

The Aymara TAME suffix system has a high degree of fusion, combining Tense and Evidentiality with number and person in obligatory suffixes. The temporal distinctions are not as clearly outlined as in Karitiana. *FIRSTHAND* pairs with general, near remote, and far remote *PAST*, and there are suffixes for *INFERRED FUTURE* and remote *INFERRED FUTURE*. The general *PAST* suffix codes every time up to the present moment, so includes past and present meaning, and stands in opposition to the *FUTURE* suffix (Hardman 2001, 104). The future is seen as not visible, behind the speaker, and the time flows from behind over the shoulder to the front, the visible past. The moment it is visible, the present sets in (ibid. 112). It may seem that Aymara is a language with a three-way distinction because it also features markers for “near remote” and “far remote”, but these do not represent a general past meaning and rather serve as specifiers in remoteness.

The Wari’ Tense system shows a similar high degree of fusion. *PRESENT/PAST* and *FUTURE* are fused with *REALIS* (but not *IRREALIS*) and in certain cases with person, number, and gender. *PAST/PRESENT REALIS* morphemes also encode person and number in a paradigm of verbal enclitics in simple sentences, and *PAST/PRESENT REALIS* also encodes gender in inflectional morphemes in different kinds of clauses. *FUTURE* is also fused with *REALIS* in both verbal enclitics and inflectional morphemes (Everett & Kern 1997, 310). So Wari’ distinguishes two sets of Tenses: *PAST/PRESENT* and *FUTURE*, regardless of whether they are expressed by verbal enclitics or inflectional morphemes. These Tense markers cannot be used in some constructions, but seem mostly obligatory. Wari’ features three optional sentence final particles for a remoteness distinction in the *PAST*, but not the *FUTURE*: *ne* ‘recent past’, *pane* ‘mid/ remote past’, and *pacara pane* ‘historic past’ (ibid. 311). The *FUTURE* morphemes also express *INTENTIONAL* and *Epistemic Modality*, and Everett & Kern (1997, 313) argue that because of this the preferred way to express future meaning in Wari’ is the use of a verbalized construction. Most Tenses are absolute, but not the verbalized *FUTURE* Tense construction (ibid. 314).

(3.14) Wari’ (Chapacuran; Everett & Kern 1997, 311, 312f.)

- (a) *xain ina-in xim ne.*
 hot 1SG.REA.PRS/ PST-3N night REC.PST
 ‘I had fever last night.’ (spoken the next morning)

- (b) *cao' 'ac cacama-in tara mijac.*
 eat travel 3PL.F-3N 3SG.REA.FUT pig
 'Then they will eat pig.'

3.4.3 Simple split FUTURE/ NONFUTURE

Simple split languages overtly marking FUTURE but NONFUTURE being unmarked are:

Kanoê, Apurinã, Yanasha', Chimila, Bororo, Kaingang, Timbira, Munich, Emérillon, Kamaiurá, Itonama, Nheengatú, Sateré-Mawé, and Yurakaré.

Kanoê does not mark for PRESENT or PAST morpho-syntactically, but has a FUTURE suffix *-nu* (which also encodes inchoative). The distinction between PRESENT and PAST is marked by person inflection on the auxiliary *re* with active verbs in the PRESENT, which also encodes PROGRESSIVE; for the PAST the auxiliary does not receive this extra person marking; this also expresses PERFECTIVE (Bacelar 2004, 221).

(3.15) Kanoê (unclassified; Bacelar 2004, 211, 223)

- (a) *iriri-ro n-e-re* person marking (PRESENT)
 run-CLF 3-DECL-AUX
 'He is running.' (orig. 'Está correndo.')
- (b) *iriri-ro e-re* no person marking (PAST)
 run-CLF DECL-AUX
 'He ran.' (orig. 'Correu.')
- (c) *aj (oj) vara-ô-ro-ô-to-nu k-e-re*
 1SG speak-1-CLF-3-TR-FUT NEG-DECL-AUX
 'I am not going to talk/ speak with him.' (orig. 'Eu não vou falar com ele.')

Like Kanoê, Apurinã does not have PRESENT or PAST marking, but a FUTURE morpheme *-ko*; indeed, FUTURE is the only overtly marked Tense. It usually refers to a non-immediate time in the future, whereas a clause unmarked with *-ko* refers to PRESENT, PAST, or immediate FUTURE. Interestingly, the FUTURE marker is not restricted to verbs (Facundes 2000, 513).

(3.16) Apurinã (Arawakan; Facundes 2000, 513, 410, 514)

- (a) *p-imaka-ko*
 2SG-sleep-FUT
 'You'll sleep.'
- (b) *nhi-nhipoko-ta-ko*
 1SG-eat-VBLZ-FUT
 'I will do eating.'

- (c) *apakata-ko kamōi-sawaku (...)* *apakata-ko uwā a-sa-ru*
 after-FUT summer-TEMP (...) after-FUT there 1PL-GO-3M.OBJ
 ‘Later in the summer (...) then we go there.’

Not being restricted to verbs is also a feature of the Yanéssha’ FUTURE enclitic. Near FUTURE can be marked by an auxiliary (*o’ch*), FUTURE by an enclitic (*-cha’ -ch/ -V’ch*); PRESENT and PAST are usually marked by temporal adverbs (Duff-Tripp 1997, 117). The FUTURE marker is apparently not restricted to the verb, e.g. is attached to *tsapať* ‘tomorrow’ in (3.17a). There are two suffixes which indicate ‘tomorrow, early’ (*-a’m*) and ‘late, night’ (*-a’n/ -n*), which do not qualify as Tense markers but rather specify a moment of the day. They must occur with the Aspectual suffixes for durative (*-en*) and regressive (*-err*) (ibid. 90):

- (3.17) Yanéssha’ (Arawakan; Duff-Tripp 1997, 117, 91)
 (a) *Tsapat’cha’ ahuen.*
 ‘Tomorrow I’ll go.’ (orig. ‘Mañana iré.’)
 (b) *Yerramu’en.*
 ‘We eat in the morning.’ (orig. ‘Comemos por la mañana.’)

The simple split of FUTURE/ NONFUTURE is the second most frequently occurring Tense system in the sample, only outranked by the FUTURE/ PAST distinction, but considerably more frequent than the FUTURE/ NONFUTURE binary split. It would be interesting to see how the numbers would change if one took zero markers to be overt markers as well. Most likely, several of the FUTURE/ NONFUTURE simple split systems would then become binary splits.

3.4.4 Binary split PAST/ NONPAST

Languages with PAST/ NONPAST binary split systems occur twice as frequently as FUTURE/ NONFUTURE binary split systems, but still account for only about a tenth of the SAILs in the sample. Languages with a PAST/ NONPAST binary split system are Warao, Hixkaryana, Panare, Rikbaktsa, Movima, and Matses.

The Tense system of Warao is heavily fused with Aspect in the PAST. It has two markers for PAST: suffix *-i* which also encodes COMPLETIVE and PERFECTIVE, and *-e* for PAST and IMPERFECTIVE. The NONPAST suffix *-te* marks PRESENT and FUTURE Tense (inherently IMPERFECTIVE) (Romero-Figueroa 1997, 95). The NONPAST suffix also carries an element of uncertainty (Romero-Figueroa 1997, 96).

- (3.18) Warao (unclassified; Romero-Figueroa 2003, 99, 92)
 (a) *naba -ya ine naru-n-a-e*
 river-ALL I go-SG-PNC-PST
 ‘I went to the river for an instant.’

- (b) *noboto-ma saba rihaba kona-te*
 child-PL for sweet bring-NPST
 '(He) brings candies for the children.'

Hixkaryana Tense suffixes mark a general NONPAST and immediate, recent, and distant PAST (Desmond C. Derbyshire 1985, 136). The NONPAST can refer to PRESENT, FUTURE, and universal Tense, though Derbyshire does not specify what he means with universal. A distinct suffix set for NONPAST DUBITATIVE occurs usually in interrogative clauses (Derbyshire 1985, 138).

(3.19) Hixkaryana (Cariban; Derbyshire 1985, 190-191., 190, 196, 197)

- (a) *ti-nyahma-yatxhe*
 1INCL.SBJ.3OBJ-supply.with.food-NPST.COL
 'We (incl) will supply them with food.'
- (b) *ti-ake-noh-no*
 1INCL.SBJ.3OBJ-burn-CAUS-IMM.PST
 'We (incl) burned it.'
- (c) *ni-niki-yako*
 3SBJ-go.to.sleep- REC.PST.COMPL
 'He went to sleep.'
- (d) *w-ama-ye*
 1SBJ.3OBJ-fell-DIST.PST.COMPL
 'I felled it.'

Hixkaryana differs from Warao in that it has no general PAST marker, but that the speaker has to choose between three remoteness degrees. In this case, the opposition is not visible as such, as there is no opposition between two markers, but between two features (the same applies to Matses). Nevertheless, Hixkaryana has a basic PAST/ NONPAST system.

3.4.5 Three-way split Tense systems

Additionally to a two-way split, discussed in the section above, some SAILs show a three-way split, i.e. overt marking of all three absolute Tenses with different markers. None of these are unmarked, but see e.g. Tiriyó for a possibly zero marked PRESENT PERFECTIVE marker below. Languages systematically marking three absolute Tenses (PRESENT, FUTURE, PAST, regardless of remoteness degrees) are Tariana, Tiriyó, Sabanê, Desano, Embera, Leko, Aguaruna, Cubeo, and Mamaindê. The systems of the first three languages are illustrated below.

Tariana encodes PRESENT, recent PAST, remote PAST, and FUTURE with verbal enclitics. PRESENT and PAST Tenses are fused with Evidentiality and all Tenses are obligatory except for when context governs the temporal interpretation or when the temporal frame has already been established (Aikhenvald 2003, 289). FUTURE Tense is not fused with Evidentiality, but there is a distinction of Epistemic Modality in the first person. Two verbal enclitics can be

used to express FUTURE: *-de*, which is only applicable with the first person and adds a sense of definiteness, and *-mhade*, which expresses a less definite FUTURE in the first person but makes no definiteness distinction for the other persons (ibid. 320).

(3.20) Tariana (Arawakan; Aikhenvald, 290, 287, 302, 320, 321)

- (a) *wa-whe-ri* *di-ñami-ka-naka*
 1PL-grandparent-M 3SG.N.F-die-DECL-PRS.VIS
 ‘Our grandfather is dying.’
- (b) *tfinu niwhã-ka* *di-na*
 dog 3SG.N.F+bite-REC.PST.VIS 2SG.N.F-OBJ
 ‘The dog bit him (we have seen it).’
- (c) *wa-whe-ri-miki-ri* *hiwiyasi-ne* *di-ñami-nhina*
 1PL-grandparent-M-NMLZ.PST-M poison-INST 3SG.N.F-die-SI.REM.PST
 ‘Our late grandfather had died through poison.’
- (d) *kasina-way-nuku* *nu-wheni-de* *nuhua nu-kale-nuku*
 now-EXTRAL-TOP.NSBJ 1SG-revenge-FUT.CERT I 1SG-heart-TOP.NSBJ
 ‘Now I will take revenge for my heart (which was devoured by an evil spirit).’
- (e) *ha-ma-pe* *pedalie-ma-pe* *na-yami-mhade-thuy* *na:*
 DEM.INA-CLF.F-PL old-CLF.F-PL 3PL-die-FUT.NCERT-all 3PL+GO
 ‘These old women will all die.’

Tiriyó has a basic distinction between factual and non-factual mood. Tense marking of PRESENT, PAST, and FUTURE only occurs in the factual mood which is further split into IMPERFECTIVE/ PERFECTIVE marking which in turn is again split into collective/ non-collective. PRESENT IMPERFECTIVE and FUTURE IMPERFECTIVE forms also encode CERTAINTY/ UNCERTAINTY (Meira 1999, 297ff.), although CERTAINTY and DUBITATIVE markers can be added as well. The only unmarked categories are the PRESENT PERFECTIVE collective and non-collective, which suggests that this is a zero marker.

(3.21) Tiriyó (Cariban; Meira 1999, 304, 307, 304)

- (a) *konopo n-eh-ta-n kokoro*
 rain 3S_A-come-FUT.IPFV-DUB tomorrow
 ‘It will rain tomorrow.’
- (b) *suurinam_po w-ei-ne,*
 Surinam_LOC 1S_A-COP-PST.PFV
wĩ-tën-ne serë_pëë
 1S_A-go-PST.PFV 3INA.PROX_ABL
 ‘I was in Surinam (then, long ago), I went there from here.’
- (c) *ji-npo kĩ-rĩ-ta-e*
 1-on.back 1+2A-do-FUT.IPFV-CERT
 ‘I will put you on my back.’

Sabanê obligatorily marks Tense and Evidentiality; PRESENT, PAST, and FUTURE also encode sensory and neutral Evidentiality and inferential and INFERRED neutral only occur in the PAST.

(3.22) Sabanê (Nambikwaran; Araujo 2004, 145, 142, 145)

- (a) *towali ilul-i-dana*
 1.SBJ to.eat-VS-PRS.E
 'I eat.'
- (b) *wayulu-mi ip-i-datinan*
 dog-REF to.run-VS-PST.E
 'The dog ran.'
- (c) *ilul-i-tapanal-i*
 to.eat-VS-FUT.N-ASSR
 'S/he will eat.'

All languages with an overtly marked three-way Tense split show a high degree of fusion with Aspect, Modality, or Evidentiality and this fusion is possibly the reason for the richness of these Tense systems.

3.4.6 FUTURE and PAST marking

A special Tense system is the overt marking of PAST and FUTURE, but not PRESENT. This is not the same as the two-way splits discussed above, as there is no binary opposition between two features/ markers, but it is not a three-way split either, as one category, the PRESENT, is unmarked. Instead, FUTURE and PAST both are marked distinct from each other and PRESENT is unmarked. This is the largest Tense system in the sample, marked by 22 SAILS:

Paresi, Jarawara, Awa Pit, Ika, Miraña, Tehuelche, Dâw, Hup, Wichí, Mosetén, Shipibo-Konibo, Yaminahua, Huallaga, Cuzco and Imbabura Quechua, Cavineña, Karo, Mekens, Cocama-Cocamilla, Tapiete, Kwaza, and Puinave.

This group is not homogeneous: some of the languages in fact only have marginal Tense marking that is an additional means of locating the verbal action in time, but not obligatory. Other languages do have obligatory Tense marking and are indeed quite close to mark three Tenses, thus almost having a three-way split. An example is Awa Pit where the absence of FUTURE or PAST markers encodes PRESENT, but only in constructions with a negative marker *ki*, and after Modal inflection of necessity. Additionally, Tense marking exists only with finite verbs, leaving subordinate clauses unmarked for Tense (Curnow 1997). Awa Pit therefore may have a three-way Tense split with a PRESENT zero marker, but unless more is known about obligatoriness of Tense marking, it remains in the FUTURE/ PAST classification. The following paragraph illustrates the FUTURE/ PAST Tense system in Jarawara as a typical example.

Jarawara overtly marks recent, distant, and far PAST, and FUTURE, but not PRESENT. All overtly marked Tenses code for gender (feminine/ masculine) and the recent, distant, and far PAST verbal suffixes are fused with Evidentiality (eyewitness/ non-eyewitness). Marking of Evidentiality is obligatory, so PAST Tense is automatically marked as well, but in general, Tense marking is not obligatory. FUTURE can be marked by the FUTURE suffix for prediction or a high degree of certainty, and an INTENTIONAL suffix which encodes planning or intention; together with a 1st incl subject it has the HORTATORY meaning ‘let us’. The FUTURE suffix is treated by Dixon (2004) as a Modal suffix, because it fills the Modal suffix slot on the verb, although its primary function is location in time and not Modality. This may point towards a Modal origin of the FUTURE suffix.

(3.23) Jarawara (Arawan; Dixon 2004, 208, 212, 213)

- (a) *o-kabe-ri-ka*
 1SG-eat-REC.PST.EYE.M-DECL.M
 ‘I ate it’
- (b) *amo o-na-habone o-ke*
 sleep 1SG-AUX-INT.F 1SG-DECL.F
 ‘I intend to/need to sleep.’
- (c) *amo o-na-habana o-ke*
 sleep 1SG-AUX-FUT 1SG-DECL.F
 ‘I will sleep.’

Since Tense in Jarawara is not obligatorily marked, there is no direct opposition between PAST and FUTURE marking and PRESENT. The PRESENT is not a zero marker in this case, but may well be on the border of becoming one, just like the Awa Pit PRESENT which already is a zero marker in certain constructions.

3.5 MAXIMUM TENSE

According to the questionnaire, the maximally possible number of Tense categories marked is five. A language with value five in Tense has markers for PRESENT, FUTURE, PAST, and remoteness degrees in both FUTURE and PAST. There are two languages in the sample that have value five: Cubeo (Tucanoan) and Embera (Chocoan). The Cubeo TAME system is characterized by a high degree of fusion, whereas Embera TAME marking tends to be analytic. For a more detailed sketch of Embera see section 3.6.8 below. In Cubeo, PRESENT, recent PAST, Evidentiality, and certain non-TAME categories (gender, person, number, and animacy) are expressed in paradigms of cumulatives. The verb status determines which Tense applies: recent PAST with dynamic verbs, PRESENT with stative verbs. In (3.24a) the suffix *-bi* marks recent PAST, FIRSTHAND information, and third person singular masculine, but if it were a stative verb, the clause Tense would be PRESENT instead. In example (3.24b) the suffix *-wi/* marks PRESENT (durative), FIRSTHAND information, first/second/ third person, and inanimacy:

(3.24) Cubeo (Tucanoan; Morse & Maxwell 1999, 40, 41)

- (a) *di-bi*
 go-3SG.M.REC.PST
 'He went recently.' (orig. 'Él fue recientemente.')
- (b) *oko da-I-wi.f.*
 water come-STV-N3
 'It is raining.' (orig. 'Está lloviendo.')

A formally different paradigm exists for non-recent PAST and PRESENT HABITUAL which are cumulatives with gender, number, person, and animacy, and all forms automatically state FIRSTHAND information (whereas in PRESENT durative and recent PAST paradigm a choice of Evidentiality can be made) and indicative. The temporal frame is again determined by whether the verb is stative (PRESENT HABITUAL) or durative (non-recent PAST). A historical PAST is marked by the suffix *-Rexa*, which denotes a time at least 20 years in the past, but sometimes only 10 years. This suffix is followed by the appropriate non-recent PAST suffix (Morse & Maxwell 1999, 46f.). PRESENT is also expressed by a paradigm of NONFUTURE nominalizer suffixes including information about animacy, number, gender, and stative/dynamic (ibid. 47). Futurity is expressed either by a set of verbal suffixes for near FUTURE, including information about number and gender (ibid. 49), or by verbal suffixes for indefinite FUTURE, including information about person, number, animacy, and DUBITATIVE/CERTAINTY. Whereas the near FUTURE suffix is simply added to the verbal stem, the indefinite FUTURE is more complicated. The verbal stem is followed by two suffixes: first, by a nominalizing suffix which also expresses indefinite FUTURE and marks for gender, number, animacy and DUBITATIVE; this is followed by another suffix of a set marking person, number, animacy, DUBITATIVE, and indefinite FUTURE (ibid. 51). Thus, Cubeo grammatically marks PRESENT, non-recent PAST, recent PAST, near FUTURE, and indefinite FUTURE in combination with various Evidential meanings, HABITUAL, Epistemic Modality, and non-TAME related information such as number, person, and animacy.

3.6 REMOTENESS DEGREES

3.6.1 Overview

Out of the 63 languages in the sample, 29 grammatically mark remoteness distinctions in Tense. Remoteness distinctions refer to a temporal distance between S and E, i.e. the interval between the points of speech and the point of event:

S ----->E

The temporal distances as well as the number of distinctions vary from language to language, but the SAILS can be arranged into three groups according to in which Tense they encode remoteness degrees:

1. Remoteness in the FUTURE, but not PAST
2. Remoteness in the PAST, but not FUTURE
3. Remoteness in both PAST and FUTURE

Additionally, they can be grouped into how many remoteness degrees they encode, ranging from one to more than four, and whether they have the same amount of distinctions in PAST and FUTURE (which does not necessarily equal the same distances) (see table 3.4).

SAILS in the sample with remoteness distinctions range from having only one remoteness degree in either PAST or FUTURE to having several in both. Although some languages are balanced and show an equal number of remoteness degrees in both Tenses the trend seems to be in favor of having more PAST remoteness distinctions than FUTURE ones. The number of languages with remoteness in the PAST but not in the FUTURE is significantly higher (18:3). Most of the languages with a remoteness degree show it in the PAST. The amount ranges from one to six degrees (five degrees are missing).¹ The following sections discuss the different kinds of remoteness in detail.

Table 3.4: Tense remoteness degrees (numbers in brackets refer to: (PAST/ FUTURE))

Remoteness degrees in the...	Languages
PAST	Puinave (2/ 0), Kwaza (1/ 0), Movima (2/ 0), Leko (1/0), Tariana (2/ 0), Jarawara (3/ 0), Aymara (2/ 0), Hixkaryana (3/ 0), Tiriyo (1/0), Wari' (3/ 0), Tehuelche (2/ 0), Aguaruna (4/ 0), Wichí (6/ 0), Mamaindê (3/ 0), Matses (3/ 0), Cavineña (2/ 0), Desano (1/ 0) , Panare (3/ 0)
FUTURE	Dâw (0/ 1), Mekens (0/ 1), Yanesha' (0/ 1)
PAST and FUTURE	Tapiete (2/ 2), Miraña (2/ 1), Northern Embera (1/ 1), Hup (2/ 2), Yaminahua (4/ 1), Cubeo (3/ 2), Cocama-Cocamilla (3/ 2), Karo (2/ 1), Shipibo-Konibo (5/ 1)

3.6.2 One remoteness degree in the PAST

Languages with one remoteness degree have a morpho-syntactic marker for a PAST Tense that specifically refers to remoteness and that exists either independently from a general PAST marker that has no remoteness specification, or is the only marked PAST.

The languages with one remoteness degree all mark for PAST that is removed far away from the PRESENT: Kwaza, Leko, and Tiriyo. Kwaza has a PAST marker *-ky* that does not have specific remoteness values and a remote PAST marker *-ĩ?ĩ* that can refer to any event in the PAST “from relatively remote to the first days of mankind” (Van der Voort 2004, 390-391). This marker can occur on verbs but also on nouns, although still referring to the verb time, and may even have scope over a noun:

¹ For a language with five remoteness degrees see Yagua (Payne & Payne 1990).

(3.25) Kwaza (unclassified; Van der Voort 2004, 392, 391, 392)

- (a) *segunDafera-iʔi-'wy*
Monday-REM.PST-time
'last Monday'
- (b) *hako'ri isi-iʔi-'wy*
moon die-REM.PST-time
'At the time there was an eclipse of the moon (one year ago).'
- (c) *ja a'nu-iʔi-hy-ki atxi'txi*
already plant-REM.PST-NMLZ-DECL maize
'He planted maize already very long time ago.'
- (d) *o'kja-iʔi-nahere*
hunt-REM.PST-COL
'those hunters of the past (of olden times)'

In Tiriyo, the remote PAST is limited to events of a 'mythical' time, and is found almost exclusively in narratives (Meira 1999, 334f.). It cannot appear in contexts like 'last Monday' as in Kwaza above (3.25a).

(3.26) Tiriyo (Cariban; Meira 1999, 335)

- kaikuki t-ëntahka-e akuri_ja*
jaguar REM.PST-fool-REM.PST agouti_AG
'The agouti fooled the jaguar.'

3.6.3 Two remoteness degrees in the PAST

Languages with two remoteness degrees in the PAST are Puinave, Movima, Tariana, Aymara, Panare, Tehuelche, Cavineña, and Desano. These all have roughly a recent vs. remote PAST distinction but vary according to where a PAST is still recent and where a remote PAST begins. That this is a highly subjective measure is seen by speaker variation within one language, for example in Puinave and Cavineña: The Puinave suffixes *-ni* and *-di* with some speakers indicate an opposition between what happened the same day (*-ni*) and what happened yesterday and more removed (*-di*). This points toward a hodiernal (same day)/prehodiernal (before today) remoteness distinction for Puinave PAST (as coined by Dahl 1984, 112; see also Dahl & Velupillai 2011b), if it were not for the fact that some speakers accept *-ni* for 'yesterday' as well (Girón 2008, 277).

Cavineña also has recent and remote PAST marking. An event starts being remote when it happened around a year before the point of speech; in the interval between the point of speech and one year ago it is recent (Guillaume 2008, 166f.). In Desano, the cut-off point is somewhere around two weeks before the point of speech: "Recent past is used about events that have just occurred or which occurred during the past week or two, and remote past refers to anything before that" (Miller 1999, 69). In Tariana, recent PAST refers to events that started a few minutes back to a few days ago and remote PAST refers to events that started a

long time ago, although the cut-off point is fluent (Aikhenvald 2003, 289). Karo's recent and remote/ mythical PAST particles *co* and *kán* are similarly ambiguous to their cut-off points. The remote PAST particle appears to be used for time spans longer than 15-20 years ago (Gabas Jr. 1999, 174).

(3.27) Karo (Tupían; Gabas Jr. 1999, 175, 177)

- (a) *púŋ* *o=ʔe-t* ***co***
 shoot 1SG=AUX-IND REC.PST
 'I shot.'
- (b) *õn* *ĩriŋ macéri* ***gán***
 1SG girl CAUS-heal-IND REM.PST
 'I healed the girl (long ago).'

3.6.4 Three remoteness degrees in the PAST

Languages in the sample with three remoteness degrees in the PAST are Jarawara, Hixkaryana, Wari', Mamaindê, and Matses. Jarawara, Mamaindê, and Matses mark PAST in cumulative morphemes with Evidential values; these will be illustrated first.

Jarawara has three marked remoteness distinctions in the PAST and all of them combine in cumulative morphemes with eyewitness and non-eyewitness (and gender). The remoteness is divided into immediate, recent, and far. Immediate PAST refers to a moment ago to a few months ago, recent to a few months to a few years ago, and far to many years ago, although these time spans are not absolute values and the markers are relative to each other rather than fixed on the time line (Dixon 2004, 207). The PAST eyewitness and PAST non-eyewitness forms do not match completely: immediate PAST non-eyewitness overlaps with recent PAST eyewitness (Dixon 2004, 208). This is possibly the result from a fairly recent introduction of the immediate PAST markers into Jarawara and that recent PAST had originally covered the immediate PAST time span (Dixon 2004, 209).

(3.28) Jarawara (Arawan; Dixon 2004, 203, 208, 219)

- (a) *owa* *na-tafi-are-ka*
 1SG.OBJ CAUS-wake-IMM.PST.EYE.M-DECL.M
 'It did waken me.'
- (b) *o-kabe-ri-ka*
 1SG-eat-REC.PST.EYE.M-DECL.M
 'I ate it'
- (c) *ee* *ta.tama* *to-he-hemete-mone*
 1INCL REDUP.be.many AWAY-AUX-FAR.PST.NEYE.F-REPO.F
 'It is said that we used to be many (i.e. the tribe used to be bigger than it is now).'

Mamaindê marks three remoteness distinctions in the PAST: non-distant, subdivided into recent and intermediate, and distant. All PAST Tenses are cumulative morphemes with

Evidentiality. The non-distant PAST indicates recent PAST which applies to events of the same day (including the previous night). Intermediate PAST covers anything from yesterday to decades ago until the time when the distant PAST sets in. The distant PAST refers to the early childhood of the speaker and before. The non-distant PAST VISUAL suffix *-nān* on its own refers to recent PAST, but when it combines with the suffix *-let* it means intermediate PAST. The distant past suffix *-hīn?* also codes VISUAL Evidentiality (Eberhard 2009, 445ff.). There is another set of Tense-Evidential suffixes that codes NON-VISUAL recent PAST, but the forms for intermediate and distant PAST are homophonous (Eberhard 2009, 457). The three remoteness distinctions are also made in INFERRED and SECONDHAND Evidentiality (ibid. 458, 469).

(3.29) Mamaindê (Nambikwaran; Eberhard 2009, 447, 448)

- (a) *wa-ta?lohni nahohnto? het-Ø-nān-wa*
 2POSS-old.woman much angry-3SBJ-NDIST.PST.VIS-DECL
 ‘Your mother was extremely angry (earlier today).’
- (b) *kanahale?i-tu <esola>-t^hīn*
 yesterday-FNS school-NCL.HOUSE
ka?jain?-je?-le-a-nān-wa
 write-EMP-INTER.PST.VIS-1SG-NDIST.PST.VIS-DECL
 ‘I wrote/ studied in school yesterday.’
- (c) *nahana-sihati-ijah-tu nūsa-jahon-nā?ā su?ton-Ø-hīn?-wa*
 time-ANC-DEM-FNS 1PL.POSS-old.man-PL not.know-3SBJ-DIST.PST.VIS-DECL
 ‘In ancient time, our ancestors didn’t know (that).’

Matses codes recent, distant, and remote PAST. Recent PAST FIRSTHAND refers to an action back until a month ago, distant roughly from a month to 50 years ago, and remote from 50 to 100 years ago. The remote PAST is used by old people to refer to their childhood (Fleck 2003, 399). Recent PAST INFERRED refers to an action back until a months ago, distant from about a month ago back to the speaker’s infancy, and remote to anything before distant (Fleck 2003, 406). Recent and distant PAST ASSUMED cover roughly the same time spans as the recent and distant PAST FIRSTHAND and INFERRED (Fleck 2003, 417).

Hixkaryana marks three remoteness distinctions in the PAST: Immediate, recent, and distant. The immediate PAST refers to the same day or the previous night, the recent PAST to the day before to several months ago, and distant PAST refers to all actions that are not covered by the recent PAST any more (Derbyshire 1979, 138–139). As Dahl (1984, 109) notes, the Hixkaryana Tense remoteness system is strikingly similar to that of Kamba, a Bantu language.

(3.30) Hixkaryana (Cariban; Derbyshire 1979, 138)

- (a) *oy-otaha-txowī*
 3SBJ.2OBJ-hit-IMM.PST.COL
 ‘They hit you.’

- (b) *r-onye-yatxoko*
 3SBJ.1OBJ-see-REC.PST.PFV.COL
 ‘They saw me.’
- (c) *w-ama-ye*
 1S.EXCL.OBJ-fell-DIST.PST.PFV.NCOL
 ‘I felled it (tree).’

Wari’ features three optional sentence final particles for remoteness distinctions in the past, but not the future: *ne* ‘recent’, *pane* ‘mid/ remote’, and *pacara pane* ‘historic’. The recent PAST refers to the last 12 hours, the mid/ remote PAST to an action of the day before back to many years ago, and historic PAST to an unspecified PAST before mid/ remote (Everett & Kern 1997, 311).

3.6.5 Four remoteness degrees in the PAST

Only one language in the sample morpho-syntactically marks four remoteness degrees in the PAST: Aguaruna. It marks recent, intermediate, distant, and remote PAST. Their values have to be seen in relation to each other: the recent PAST suffix *-ma* refers to earlier the same day to a few days ago and the intermediate PAST suffix *-maia* refers to a time longer ago than recent PAST. It is probably a combination of recent past suffix *-ma* and remote past suffix *-ia*. The distant PAST suffix *-amaia* is further removed from the PRESENT than intermediate PAST but closer to it that the remote PAST suffix *-ia* which refers to a time many years ago. The intermediate and distant PAST suffixes are probably grammaticalized from the recent and remote PAST suffixes, but their usage points toward independent markers. Recent and distant PAST also encode FIRSTHAND information. Aguaruna additionally has a narrative PAST suffix *-haku*, but this covers roughly the same time span as the remote PAST marker (Overall 2007, 336–345).

(3.31) Aguaruna (Jivaroan; Overall 2007, 336, 338, 343, 344)

- (a) *ti-ma-ha-I* recent
 say+LOAF-REC.PST-1SG-DECL
 ‘I said.’
- (b) [*nu-na muunta auhumatu-inu aha-maia=nu-na*] intermediate
 ANA-ACC elder tell-NR COP.PST-INTER.PST.3=ANA-ACC
wi-ʃa ta-ha-I
 1SG-ADD say+IPFV-1SG-DECL
 ‘that (story) which the elders told, I also tell it (now)’
- (c) *ta-aw-amaia-ha-I* distant
 come-HIAF-DIST.PAST-1SG-DECL
 ‘I arrived.’

- (d) *tuna-na-fa kahama-inu a-ia-ha-I* remote
 waterfall-ACC-ADD dream-NR COP-REM.PST-1SG-DECL
 'I was dreaming of a waterfall too.' (i.e. 'I saw a vision of spirits.')

3.6.6 Six remoteness degrees in the PAST

Wichí is the only language in the sample that has six remoteness degrees in the PAST, which is also the highest number of remoteness degrees in both PAST and FUTURE in the sample. Wichí has shows several clitics for PAST Tense whose remoteness values range from immediate to remote. Their order relative to each other is shown below:

remote past, fairy tales>historical, mythical>years ago>yesterday>present
saxi p'ante te naxi mati, ne
 (from Terraza 2009, 161-162)

The clitic *saxi* is not productive any more but still recognized by speakers for a remote PAST. It is not clear whether it is indeed more remote than *p'ante* which is employed for narratives of historical and mythical context. The clitic *te* refers to a past of several years ago. There are three markers that refer to a recent PAST: *naxi* is used for events of the day before or at least 24 hours ago, *ne* refers to an immediate PAST, and *mati* refers to an immediate PAST on the morning of the same day (Terraza 2009, 162), although *-mati* also occurs with the translation of 'last night'. It should be noted that all productive PAST Tense clitics can also occur on nouns.

(3.32) Wichí (Matacoan; Terraza 2009, 162,189)

- (a) *ha-la-nahyt-ne yelatax*
 INTRG-2-tether-PST horse
 'Did you tether the horse (just now)?' (orig. '¿Ataste el caballo (recién)?')
- (b) *lam ta-yotsan-mat-nu(y)-a*
 PR.3SG 3-ask-PST-OBJ1-APPL
 'He asked me (just a while ago).' (orig. 'Él me preguntó.' (hace un rato))
- (c) *sinox y-uk^wax-naxi to Ø-i-hi tahy*
 dog 3-bite-PST SUB 3-be-LOC forest
 'The dog bit him when he was in the forest.' (orig. 'El perro lo mordió cuando estaba en el monte.')
- (d) *n-i-(y)ex-te n-qatela wet*
 1-be-APPL-PST 1POSS-grandmother house
 'I lived in the house of my grandmother (with her).' (orig. 'Yi viví en la casa de mi abuela (con ella).')

- (e) *nixutas-p'ante* *i-lon-hen* *wik'i*
 soldiers-PST 3-kill-PL people
 'The soldiers (of that age) kill the people.' (orig. 'Los soldados (de esa época) matan a la gente.')

3.6.7 Remoteness degrees in the FUTURE

There are only three languages in the sample that show remoteness degrees in the FUTURE, but not in the PAST: Dâw, Mekens, and Yanasha'. Dâw indicates absolute FUTURE, immediate FUTURE, and 'strategic' FUTURE. The immediate FUTURE markers are suffix *-?əj* and particle *nã?* (Martins 2004, 283)) which do not seem to have any difference in meaning. They express an interval in time that happens immediately after the point of speech. According to context, they can also express a directive (let us...) or an action that has just been initiated.

(3.33) Dâw (Nadahup; Martins 2004, 283, 199)

- (a) *me* *nã?* *m'ũg* *wujx* *ʃúk-əj*
 1PL.HORT FUT here arrive hunt-IMM.FUT
 'We are going to arrive and start hunting right away.' (orig. 'Vamos chegar e començar a caçar daqui mesmo.')

- (b) *hid* *ʃúk-əj*
 3PL hunt-IMM.FUT
 'They go hunting now.' (orig. 'Eles já vão caçar.')

Mekens has one particle *kot* to signal immediate FUTURE. Other particles signal a general FUTURE and IRREALIS FUTURE. Mekens possibly has more than one remoteness degree in the FUTURE (Galucio p.c.).

(3.34) Mekens (Tupían; Galucio 2001, 151)

- o-er-a* *kot*
 1SG-sleep-THEM IMM.FUT
 'I will sleep.'

The status of Yanasha' marking remoteness in the FUTURE remains doubtful. A normal FUTURE is marked by the enclitic *=cha'* and proximate FUTURE by the auxiliary *o'ch* (Duff-Tripp 1997, 117), but examples are scarce and a discussion in Duff-Tripp (1997) is completely missing.

It is interesting that, firstly, only three languages do have remoteness in the FUTURE but not PAST, and secondly, that this is immediate, i.e. following right after the point of speech. Dâw is of the Nadahup family of which there is one other language in the sample: Hup, which shows remoteness distinctions in both PAST and FUTURE. A search of Nadahup in Martins & Martins (1999) has yielded no information about FUTURE in other Nadahup languages. Geographically, Dâw is close to Hup and Desano, but Desano has remoteness only in the PAST. No currently known relationship, either geographically or genetically, can

therefore explain the Dâw Tense remoteness system. The same holds for Mekens, which is of the Tupían family. The other more Tupían languages in the sample do not show a similar remoteness system.

3.6.8 Remoteness in the PAST and FUTURE

Languages with morpho-syntactically marked remoteness distinctions in both PAST and FUTURE either have the same amount of degrees in both or more in the PAST, but never more in the FUTURE. Tapiete, Northern Embera, and Hup have a balanced remoteness system: Tapiete codes immediate and distant PAST and FUTURE. The immediate FUTURE is indicated by two suffixes, *-po* and *-pota*, the latter additionally encoding desire (González 2005, 155).² The distant FUTURE suffix *-kwi* can be combined with any of the immediate FUTURE suffixes for an even more distant FUTURE. Immediate and distant PAST are also marked by suffixes (González 2005, 156). There is no indication of which time spans either immediate and distant FUTURE or PAST inhabit.

Northern Embera marks immediate FUTURE and immediate PAST. The two examples given in Mortensen (1999) point towards usage of these markers referring to actions very near to the point of speech, although this cannot be claimed with certainty. A study of the texts in appendices A and B in Mortensen (1999) did not yield any further occurrence of these markers.

(3.35) Northern Embera (Chocoan; Mortensen 1999, 72)

- (a) *wã-podo-a*
 go-IMM.FUT-DECL
 'He is about to leave.'
- (b) *wã-toko-a*
 go-IMM.PAST-DECL
 'He just left.'

These immediate PAST and FUTURE markers also seem to be restricted to Northern Embera, as in the Embera grammar by Aguirre Licht (1999) about the Chamí variant (Southern Embera) no Tense remoteness degrees are mentioned at all. It is worth mentioning that the immediate PAST suffix is homophonous with the verb *toko* 'to run' (cf. Mortensen 1999, 164, examples 74 and 76).

Hup has one particle for temporal proximity and two particles for distant FUTURE and PAST respectively. It is important to realize that all the remoteness particles refer to a contrast of Tense with regard to the PRESENT; they are therefore absolute Tenses. The proximate particle *páh* refers to actions immediately before or after the point of speech and even simultaneous actions. It cannot refer to a distant Tense (Epps 2008, 600). It can co-

² According to Jensen (1998), descendant of the proto-form **potár* occur commonly in Guaraní as FUTURE or DESIDERATIVE (see also Mueller 2013).

occur with the general FUTURE marker (cf. *ibid.* 602, example 47). The distant past particle *j'ám* is common in narratives, but can also refer to the time the speaker heard the story (*ibid.* 604). The distant FUTURE particle *tán* exists synchronically parallel as an independent adverbial and can also indicate a simple FUTURE when the usual FUTURE markers are ungrammatical (*ibid.* 606). It can co-occur with the general FUTURE markers as well. Thus, although the remoteness markers in Hup are not restricted to locating an action on the time line, they do mark actions in contrast to other actions and can therefore be considered Tense markers. The time spans are not further defined. The same holds true for Miraña, which has an immediate FUTURE and recent and remote PAST enclitics (Thiesen 1996, 97–98) whose temporal distinctions are not clear.

Shipibo-Konibo is the language with probably the most remoteness degrees of the sample, together with Wichí. The PAST degrees include: a PAST suffix of the same day *-wan*, a PAST suffix for the day or a few days before *-ibat ~ ibá*, a PAST suffix for some months/ years ago *-yantán*, a PAST suffix for about nine months to three years ago *-rabe*, a distant PAST suffix for many years ago *-kati(t)*, and a remote PAST suffix *-ni* (Valenzuela 2003, 284–285). The suffixes *-yantán* and *-rabe* roughly cover the same time spans, but the former additionally codes IMPERFECTIVE and the latter PERFECTIVE. The only clearly temporally defined Tense is the PAST suffix *-wan* for an action of the same day; the other suffixes are relative to the speaker and each other and also overlap.

(3.36) Shipibo-Konibo (Panoan; Valenzuela 2003, 285, 286, 287, 288)

- (a) *rama bariapan-ra ka-wan-ke*
 now at.noon-DEV go-PST-COMPL
 'He left today at noon.'
- (b) *ikaxbi bakisha e-n maxaman kawáti*
 but one.day.away.DEV 1-ERG manxaman kawáti
ao-bo icha oin-iba-ke
 k.tree-PL much.ABS see-PST-COMPL
 'But yesterday I saw a lot of manxaman kawáti trees.'
- (c) *e-n-ra oin-yantan-ke rabé bene-ya yoxan*
 1-ERG-DEV see-PST-COMPL two husband-PROP adult.woman.ABS
 'A few years ago I met a woman who had two husbands...'
- (d) *pekáo kirika wisha-rabe-ke*
 after letter.ABS write-PST.IPFV-COMPL
 'Last summer after the morning meal my brother (usually) wrote letters.'
- (e) *moa-tian nawa-baon chop a no-n bi-ama-katit-ai*
 already-TEMP mestizo-PL.GEN cloth 1PL-ERG get-NEG-PST-INCP
 'Long time ago, we did not buy pieces of cloth in the store.'

- (f) *moa-tian-ronki i-pao-ni-ke jema-bo yama*
 already-TEMP-HEA do.INTR-HAB-REM.PST-COMPL village-PL.ABS exist.not
 'It is said that in remote times there were no villages.'

So Shipibo-Konibo has six morphemes for remoteness distinctions in the PAST, although two of them code the same time span, but differ in their Aspectual value.

Karo has two remoteness distinctions in the PAST and one in the FUTURE. Karo's recent and remote/ mythical PAST particles *co* and *kán* are ambiguous with respect to their cut-off points. The remote PAST particle appears to be used for time spans longer than 15-20 years ago (Gabas Jr. 1999, 174). Interestingly, both particles can combine to mean 'long ago', although the same meaning is already covered by the remote PAST particle. FUTURE in Karo can be expressed by the particles *iga* and *yat*, which do not specify any remoteness, and the auxiliary *kap*, which refers to immediate FUTURE. *Iga* is used exclusively in interrogative, negative clauses (Gabas 1999, 183). The particle *yat* can also occur on nouns.

(3.37) Karo (Tupían; Gabas Jr. 1999, 175, 177, 179)

- (a) *púŋ o=ʔe-t co*
 shoot 1SG=AUX-IND REC.PST
 'I shot.'
- (b) *õn ïriŋ macéri gán*
 1SG girl CAUS-heal-IND REM.PST
 'I healed the girl (long ago).'
- (c) *õn o=pi-t møy mãm co kán*
 1SG 1SG=perforate-IND long X REC.PST REM.PST
 'I took the vaccine long ago.'
- (d) *o=wíy-a o=kap-t*
 1SG=leave-GER 1SG=AUX.FUT-IND
 'I am going to leave.'

Yaminahua PAST markers are partly cumulative morphemes with PERFECTIVE/ PROGRESSIVE Aspect: PERFECTIVE PAST of same day suffix *-a* and IMPERFECTIVE PAST of same day suffix *-faiyamea* are for the same time span. For some speakers, *-faiyamea* indicates PAST of yesterday without Aspectual meaning, and *-ita* a PAST a few days to two weeks ago. For some speakers, *-ita* signals a PERFECTIVE PAST of yesterday or a few days ago. Remote PERFECTIVE PAST suffix *-ti* codes actions a week up to several years ago, as does the remote PAST PROGRESSIVE suffix *-yamea*. For some speakers, there is no aspectual distinction between the two. There is no Aspectual distinction in the far PAST suffix *-ni* (Faust & Loos 2002, 119–122) which also occurs as a remote PAST suffix in Shipibo-Konibo. The PROGRESSIVE FUTURE suffix *-fainaka* codes an action the day after the point of speech and probably consists of the suffixes *-fai* and *-naka*, meaning 'a day from now' and FUTURE PROGRESSIVE respectively (ibid. 123). The suffix

-fai probably is also a part of the PAST of yesterday suffix *-faiyamea*, together with the remote PAST PROGRESSIVE suffix *-yamea* (ibid. 120).

Cubeo and Cocama-Cocamilla are similar in that they both mark three remoteness distinctions in the PAST and two in the FUTURE. Cubeo shows a high degree of fusion: near FUTURE morphemes are cumulatives with gender and number; remote FUTURE morphemes are cumulatives with gender, number, animacy, and doubt. The recent and non-recent PAST suffixes are cumulatives with person, number, and some also with gender and animacy, while at the same time being ambiguous between PAST and PRESENT (either durative or HABITUAL) according to whether they are affixed to a dynamic or stative stem. Recent PAST refers to an action that happened within a week before and non-recent PAST to a time before that, although non-recent can also mean a few days before, being in the sphere of recent PAST. Anything before non-recent PAST is marked with the historic PAST suffix *-RExa* which is added to the non-recent PAST suffix (Morse & Maxwell 1999, 38ff.). Cocama-Cocamilla codes immediate and remote FUTURE and immediate, mediate, and remote PAST by clitics (Vallejos Yopán 2010, 161). In the FUTURE Tenses, there is an additional Modal value: The immediate FUTURE enclitic also implies relative certainty, the remote FUTURE in contrast implies probability of an event and wishes and hopes. There is no such Modal value in the PAST Tenses. The immediate PAST refers to actions right before the point of speech to a few days ago, the medial PAST from the day before until the remote PAST sets in for events long ago (childhood and even mythical times). The events range from being vivid to less strong in the speaker's memory (ibid. 471). The choice of Tense is primarily dependent on whether the speaker's recollection of the action is still vivid or less strong, rather than on exactly when the action happened, which explains the temporal overlap of the Tense markers.

Urarina presents a special case. It does not mark general absolute Tense, but it does have an enclitic for remoteness. The enclitic *=lu* marks actions in the past that are quite removed, although this depends on speakers' perception, but also remoteness in the future (Olawsky 2006, 501). There is no general Tense marker it could contrast with, and no Tense marker referring to a more immediate time (see also section 3.3).

3.6.9 Discussion

Although at first glance quite heterogeneous, the SAILS remoteness distinctions tend to adhere to the same principles: they differ according to which Tense and how many remoteness distinctions they make, but keep well within certain boundaries. The majority codes an average of about two distinctions in the PAST and one in the FUTURE, Wichí and Shipibo-Konibo being the exceptions. Shipibo-Konibo and Wichí mark the highest number of remoteness distinctions: Shipibo-Konibo marks five in the PAST and one in the FUTURE, and Wichí six in the PAST. Shipibo-Konibo actually has six morphemes for remoteness distinctions in the PAST, but two of them code the same time spans and differ only in their Aspectual value. Strikingly, only three of the SAILS have more distinctions in the FUTURE than in the PAST: Dâw, Yanesha', and Mekens, which have one remoteness degree in the FUTURE but none in the PAST. As a rule, all languages that morpho-syntactically mark

remoteness distinctions in both Tenses have either more in the PAST or are at least balanced. A second rule is that the number of degrees in the PAST is higher than in the FUTURE: Where Wichí has as many as six PAST distinctions; the highest possible number of FUTURE distinctions is two (in Tapiete, Hup, Cubeo, and Cocama-Cocamilla). Dahl (1984, 107) already pointed out that the PAST remoteness systems are more developed than those in the FUTURE: “[I]n general, the distinctions in the past appear to be more well developed – that is, to be more numerous and well defined than those in the future”. Comrie (1985, 85) similarly notes: “it turns out that the more prolific sets of distinctions are more widespread in the past than in the future, in accord with the general tendency of languages to have a better developed past than future system; there are, however, some languages which have symmetrical systems with several oppositions of temporal distance in both past and future”. Dahl’s sample consists of 60 languages distributed over the globe, and at least his claim that the distinctions are more numerous in the PAST than in the FUTURE is supported for the SAILS sample as well; the three exceptions being Dâw, Yanasha, and Mekens. Below, a more detailed study will be given of the distances and “definedness”, in Dahl’s terms.

So far we looked at the number, but not at the properties of the distinctions. The sample suggests that the distance measures of values such as ‘immediate, recent, remote, far’ etc. are not absolute and that speakers vary in their subjectivity about when an event happened/ will happen. This also relates e.g. to the age of the speaker who may refer to any past in his own lifespan as being reasonably intermediate, whereas any time before his birth is beyond imagination, i.e. far away. The same points in time can be seen differently by another person of a different age, and so the same time can be referred to by different remoteness degrees. This is not possible, of course, when a remoteness degree is clearly and well defined, i.e. is objective and independent of subjective measures. Such degrees always refer to a certain point in time. In the SAILS, some languages do have such defined remoteness degrees and these strikingly concern time spans very close to the point of speech. Mamaindê, Puinave, Hixkaryana, Wichí, Shipibo-Konibo, Yaminahua, and probably Northern Embera and Hup have markers that refer to past events of the same day, including the previous night. These are called hodiernal ‘same day’. Puinave, Wichí, Yaminahua, and probably Shipibo-Konibo also mark past events restricted to the day/ a few days before, called hesternal (Dahl 1984, 113). Please note that this list is not exhaustive as many languages have not been researched closely enough. The restriction of markers towards definite time spans points to objective values, although there is still variation among speakers. Dahl’s (ibid. 112) generalization “[if] there are one or more distinctions or remoteness in a tense-aspect system, and reference can be made to objective time measures, one of the distinctions will be between ‘more than one day away’ and ‘not more than one day away’” (number 1270 in The Universals Archive) holds for SAILS as well. Languages with such objective measures are, however, in the minority compared to languages with subjective measures. It also seems that firstly, the fewer remoteness distinctions a language has, the more they tend to be subjective, and secondly, objective time measures seem to apply mostly to time spans located immediately around the point of

speech. A look at Wichí shows that the remoteness distinctions immediately before the point of speech are much more differentiated than those further away. Three of the six PAST remoteness markers refer to the time spans ‘same day’, ‘24 hours ago’, and even ‘morning of the same day’, whereas the other three markers cover much bigger time spans. No language makes a distinction of e.g. ‘morning of a day a hundred years ago’ (note that values like ‘day’ and ‘year’ may not apply to every culture and are results of a researcher’s interpretation).

Although most languages with remoteness distinctions in both PAST and FUTURE do have a more elaborate system in the PAST, there are three languages with a balanced system: Tapiete, Northern Embera, and Hup. Tapiete and Hup mark a PAST/FUTURE that is quite near to the point of speech and in contrast to that another that is more distant. Northern Embera marks immediacy in both Tenses. Although no definite time spans are deductible from the sources, it can be claimed that languages that have a balanced remoteness system also mirror the remoteness distinctions with the PRESENT being the pivotal point.

While marking for immediacy is quite common, some languages mark for a PAST described as mythical or historical. The actions marked by those Tenses are often outside the speakers’ personal recollection and refer to myths, fables, or creation stories. At the same time, those markers can have a special narrative value: The Wichí’ clitic *p’ante* is used for mythical/ historical PAST and at the same time indicates that the speaker is narrating a story; Hup employs the distant PAST particle *jám* for narratives. Aguaruna even has an additional suffix for narrative PAST, which covers the same time span as the general remote PAST. Tiriyo, Aymara, Wari’, and Cubeo have forms for expressing a historic PAST. Apparently speakers of languages with those historical or mythical PAST Tenses feel the need to differ between events that are not only far removed from the PRESENT but also that these events are narrated and have not been experienced by the speaker. This indication of how the information was acquired is commonly known as reported or second-hand Evidentiality and will be dealt with in chapter 6.

Some languages allow for a combination of remoteness markers, although this does not necessarily result in a different time span, i.e. is not iconic. In Tapiete, the distant FUTURE suffix together with the immediate FUTURE refers to an even more distant FUTURE, but in Karo, the combination of remote and recent PAST markers simply refers to the same time span as the remote PAST, and in Cubeo, the historic PAST suffix together with the non-recent PAST refers to non-recent PAST as well. Closer inspection suggests that in Karo and Cubeo a double marking has emphasis value rather than increasing the remoteness degree.

The study of remoteness degrees in the SAILS sample allows for the following generalizations (additionally to those already stated by Dahl (1984) and Comrie (1985) above):

- When a language has remoteness degrees in PAST and FUTURE, the number of those in the PAST is higher or identical to those of the FUTURE.

- When a language has the same amount of remoteness degrees in PAST and FUTURE, these refer to the same time spans.
- Languages differentiate between subjective and objective time measures. When a language has objective time measures, they apply to time spans close to the point of speech. The more distant a time span, the less objective it is. And, vice versa: The more recent a time span, the more defined (and objective) it is.

If one compares the languages according to their Tense system and which remoteness degrees they show, remoteness degrees occur most frequently in languages with a FUTURE/PAST system and least in PAST/NONPAST and FUTURE/NONFUTURE systems. This is not surprising, because FUTURE/PAST is the most frequently marked system in the complete sample. When we look in detail at the remoteness degrees in PAST/NONPAST and FUTURE/NONFUTURE, we see no significant differences: three languages with a PAST/NONPAST Tense system have more remoteness degrees in the PAST than in the FUTURE and none more in the FUTURE; two languages with a FUTURE/NONFUTURE system have more remoteness degrees in the PAST and one more in the FUTURE (see table 3.5). In Ultan's (1978) terms, there is no significant difference between remoteness marking in prospective and retrospective languages. This is, however, not in line with Ultan's (ibid. 92) results which show that eight retrospective languages have more remoteness degrees in the PAST than in the FUTURE contrary to zero prospective languages (in a sample of about 50 languages Ultan unfortunately does not specify further). To summarize, there is no significant correlation between Tense system and remoteness degrees.

Table 3.5: Cross-comparison of Tense systems with remoteness degrees

	No remoteness degrees	More remoteness degrees in the PAST	More remoteness degrees in the FUTURE	balanced
PAST/NONPAST (prospective)	Warao, Rikbaktsa	Hixkaryana, Matses, Movima, Panare	-	-
FUTURE/NONFUTURE (retrospective)	Karitiana, Apurinã, Chimila, Bororo, Kaingang, Timbira, Kamaiurá, Itonama, Yurakaré, Kanoê, Emérillon, Nheengatú, Sateré-Mawé	Aymara, Wari'	Yanesha'	-
Three-way split	Sabanê	Tariana, Tiriyo', Desano, Leko, Aguaruna, Mamaindê, Cubeo	-	Northern Embera
FUTURE/PAST	Paresi, Ika, Awa Pit, Mosestén, Huallaga Quechua, Cuzco Quechua, Imbabura Quechua,	Jarawara, Miraña, Tehuelche, Wichí, Shipibo-Konibo, Yaminahua, Cavineña, Karo, Cocama-Cocamilla, Kwaza, Puinave	Dâw, Mekens	Hup, Tapiete

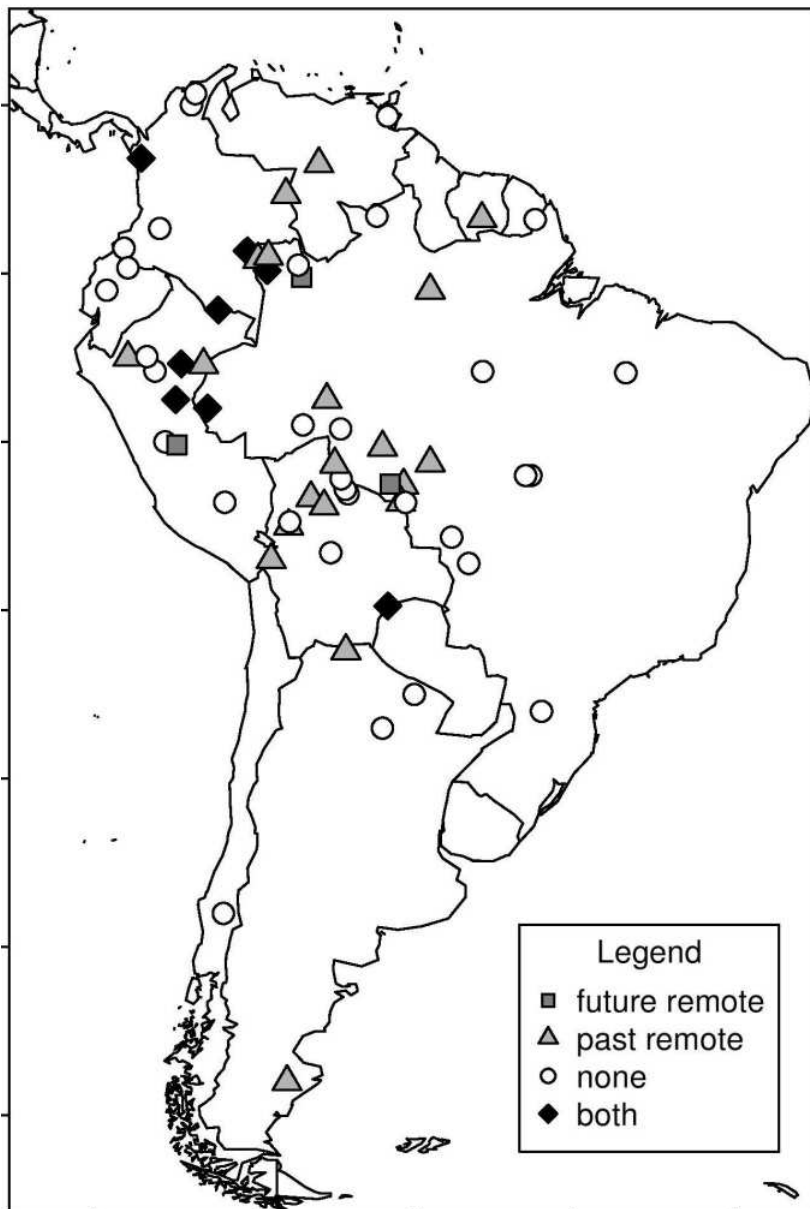
The sample reveals no clear geographical pattern as to marking of remoteness degrees (cf. map 3.1). Remoteness degree marking is spread all over the continent and families: it is marked in Tupían, Tucanoan, Arawakan, Boran, Cariban, Chocoan, Chonan, Jivaroan, Nadahup, Matacoan, Nambikwaran, Panoan, Tacanan, Tucanoan, and Tupían, but absent in Barbacoan, Chibchan, Macro-Gêan, Mapuche, Mosétenan, Paezan, Quechuan, and Yanomaman. In total, 30 languages morpho-syntactically mark at least one remoteness distinction. There is, however, evidence for remoteness distinctions being particularly pronounced in SAILS in comparison to the rest of the world. In Dahl & Velupillai's (2011b) global sample of 222 languages, only 40 languages have remoteness degrees and only two of them have four or more. These 40 languages cluster especially in South America (and Papua New-Guinea) and the two languages with the highest number of remoteness degrees marking are also found in SA (Chácobo and Yagua). Although most SAILS in the sample do not have a very high number of remoteness degrees, the overall tendency of SAILS to mark (many) remoteness degrees as found by Dahl & Velupillai (2011b) is confirmed.

Considering the distribution of languages with a high number of remoteness degrees a small cluster occurs. Whereas Wichí (six remoteness degrees) and Cubeo (5) are isolated instances, Shipibo-Konibo (6), Yaminahua (5), and Cocama-Cocamilla (5) form a geographical cluster. Shipibo-Konibo and Yaminahua are Panoan languages, but Cocama-Cocamilla is Tupían (Guaraní). Whether the latter has acquired an elaborate remoteness degree system by contact with Panoan has to be seen, but it is certainly the case that a large remoteness degree system is a Panoan characteristic.

Map 3.1 illustrates the distribution of languages in the sample with no remoteness degrees, degrees in the FUTURE or PAST, or both.

Map 3.1: Distribution of remoteness degrees in the sample

Remoteness degrees



3.7 TYPOLOGICAL DISTRIBUTION IN THE SAMPLE

The Tense marking of SAILs is not homogeneous, but allows for a rough classification into the four groups two-way split, three-way split, FUTURE/ PAST and Tenseless, as shown in the section 3.4. If more detailed information about some languages becomes available, future studies may lead to different analyses in some cases. Based on the groupings, several questions can now be looked at in more detail: what do Tense systems tell us about geographical and/ or genealogical relations? What are the implications for universals or universal tendencies in Tense? The first question will be discussed now and the latter later in this section.

Out of the 63 languages in the sample, seven do not mark any Tense, 12 mark one, 11 mark two, 15 mark three, 16 mark four, and two mark five Tenses. The most frequently marked category is FUTURE which is marked by 54 languages, followed by PAST (44), remoteness in the PAST (28), PRESENT (16), and remoteness in the FUTURE (11) (see table 3.7). The following generalizations can be made:

1. When a language morpho-syntactically marks Tense, it marks at least FUTURE (the two exceptions are Mapuche and Trumai).
2. When PRESENT is marked, PAST is always marked as well.
3. No language marks only PRESENT or PAST.
4. Remoteness in the FUTURE is always marked when FUTURE is marked; remoteness in the PAST is always marked when PAST is marked.

The results can be formed into a hierarchy of Tense marking: FUTURE marking occurs more frequently than PAST occurs more frequently than PRESENT. The rather low number of languages marking PRESENT is probably due to the exclusion of zero markers in this study; in quite a lot of those languages that do not overtly mark PRESENT it is understood as the absence of other Tense marking or simply as default. This in turn leads to the conclusion that FUTURE is understood to be furthest removed from the default. A study similar to this one was carried out by Ultan (1978, 116) who also found that FUTURE is “more marked”. However, Ultan neither presents a language list (“approximately fifty languages” (ibid. 85)) nor does he explain satisfactorily what he understands by ‘marking’: “any grammatical form (affix, particle, auxiliary, etc.)” (ibid. 87). In contrast to studies by e.g. Comrie (1985), Stassen (1997), Smith (2008), and Bybee et al. (1994) obligatoriness is not a defining feature for Tense here. It would be interesting to investigate whether the maxim that FUTURE is marked more frequently than PAST and PRESENT still holds when applied only to obligatory Tense markers.

All languages without morpho-syntactic Tense marking have other means of expressing temporal relations: by temporal lexemes and/ or context and/ or by morphemes of Aspect, Modality, or Evidentiality etc. No language lacks the means of indicating time as such, but it is surprising that out of a sample of 63 only seven SAILs do not have any morpho-syntactic Tense marking, especially with regard to the rather broad definition of morpho-syntactic

marking in this study. On the other hand, this means that Tense marking is not a universal in the sample, which was already shown with slightly different definitions of Tense by e.g. Comrie (1985) and Stassen (1997) for global samples.

Generalizations 1. and 2. can be explained by the design of the questionnaire and the occurrence of cumulative Tense markers, i.e. NONFUTURE OR NONPAST. These usually occur in systems with an overt marker and an unmarked default.

The most frequent Tense systems in the sample are FUTURE/ PAST and two-way splits; likewise, three-way splits and Tenseless systems are least frequent. Within the two-way splits, FUTURE/ NONFUTURE outnumber PAST/ NONPAST. The theoretically also possible PRESENT/ NONPRESENT split does not occur, suggesting that no (absolute) Tense marker can refer to a discontinuous time, i.e. cannot refer to non-adjacent Tenses. This has also been observed by Comrie (1985, 50): “a possible universal of tense systems: in a tense system, the time reference of each tense is a continuity. If this universal can be maintained in general, then it would exclude the possibility of discontinuous tenses”. This possible universal is thus true for the 63 SAILs in this sample as well, but is possibly not applicable to remoteness degrees. The case of the remoteness enclitic =~~h~~ in Urarina (see section 3.3) shows that one marker can refer to either remote FUTURE OR PAST, i.e. this is a discontinuous Tense marker. But the fact that this is the only remoteness marker in the sample ambiguous between PAST and FUTURE suggests that this rather an exception and that the continuity principle holds for remoteness degrees as well.

PRESENT Tense in general is least marked in the sample; it overtly shows only in three-way splits and is at best a zero marker in other instances, but mostly being ambiguous with either PAST OR FUTURE. This raises an interesting question about the presence of the PRESENT Tense in the human mind and its status as a default category, as well as the importance of one absolute Tense over another, or the perceived irregularity vs. default. It is difficult to assign importance to abstract features such as TAME, because it may depend on speaker-internal rather than language-internal rules and thus crosses the border from typology to psycholinguistics. But independent from why the PRESENT Tense is least marked, it is a fact that the SAILs show a preference towards marking FUTURE OVER PAST OVER PRESENT. When a language marks one category and is ambiguous about the other two, it is usually FUTURE that is marked and PRESENT and PAST occur in a cumulative morpheme. This may be explained by a tendency of speakers to assign FUTURE a kind of unreal perception that is most removed from normality, or default, and which therefore obtains extra marking, in contrast to the less removed and more ‘normal’ perception of PRESENT. But this may also be due to the fact that, despite the careful investigation of each marker, some or many of the FUTURE markers are Modals in disguise and should not be encoded as Tense at all. Ultan (1978), Bybee & Pagliuca (1987), and Bybee et al. (1994), besides others, showed that FUTURE markers are accompanied by Modal flavors of e.g. desire, intention, obligation, necessity, imminence, Epistemic Modality, commands etc. Although this is the case for many if not all FUTURE markers in this study these predominantly represent a temporal meaning (see the dominance parameter in section 2.3.2), and are therefore not Modals.

It has been previously suggested that FUTURE indeed is not even a Tense at all, but rather a Modal. Beside the semantic argument (see paragraph above) the one of “formal asymmetry” is brought forth: often FUTURE markers are encoded differently than PRESENT and PAST in the same language, i.e. do not form a morphologically consistent system. Ultan (1978, 91) was probably the first to observe this asymmetry of formal marking of FUTURE in Tense systems: “[f]uture tense markers may be less bound than present or past but never more so” (number 1387 in the Universals Archive). Thus, a language should not be able to have a periphrastic PAST and a FUTURE suffix, but the other way round is possible. For this study, this means that there should be no language with a more bound FUTURE than PRESENT or PAST; the scale extending from least bound to most bound:

auxiliaries > particles > clitics > affixes

In this sample, most languages with FUTURE and either PRESENT or PAST or both markers have the same formal class of marking, i.e. all are suffixes or prefixes or enclitics or particles. Suffixation occurs most frequently, probably because suffixation is in general the most frequent form of Tense marking in the sample. An example for marking both FUTURE and PAST by the same formal means (enclitics) is found in Cocama-Cocamilla:

(3.38) Cocama-Cocamilla (Tupían, Guaraní; Vallejos Yopán 2010, 476, 478)

- (a) *ay ikuachi r=ichari=ta=ikua*
 already yesterday 3SG.M=leave=1SG.M=PST
 ‘Already, yesterday, he left me.’
- (b) *r=eyu=utsu ra tsai=kira=pu*
 3SG.M=eat=FUT 3SG.M tooth=DIM=INST
 ‘He will eat with his little teeth.’

Languages with a FUTURE marker that is less bound than a PRESENT or PAST (or both) markers are rare in the sample; most languages have a balanced formal marking. An example where FUTURE is less bound is found in Mekens, where the FUTURE is marked by particles and the PAST by suffixation (PRESENT is not marked):

(3.39) Mekens (Tupían; Galucio 2001, 62, 91)

- (a) *paroray at paat ôt kîrêp masopi=bõ*
 armadillo get FUT I today night=DAT
 ‘I will hunt Armadillo tonight.’
- (b) *o-kwe-a-r=õt kîpkiba=bõ*
 1SG-climb-THEM-PST=I tree=DAT
 ‘I climbed on the tree.’

Whereas indeed in most languages with FUTURE and PRESENT and/ or PAST markers in the sample Ultan's proposed universal holds, there are a few exceptions: Miraña, Hup, Wichí, and Kanoê. For example, in Miraña FUTURE is obligatorily marked by the suffix *-i* or vowel lengthening (depending on the verb stem) (Seifart 2005, 62) and PAST by a set of optional enclitics that attach to the first clause constituent (ibid. 72). Both Tenses occur as separate systems in Miraña which may explain why FUTURE is more bound. Further research is necessary to compare the systems of Miraña, Hup, Wichí, and Kanoê.

Besides giving FUTURE preference of marking, the sample also shows that obligatory Tense marking is in general restricted to certain syntactic constructions and also relies on semantic considerations. Tense marking can be quite restricted (cf. Awa Pit, where Tense is not marked in subordinate clauses), and although in languages where Tense marking is said to be obligatory it can be dropped once the temporal frame is established (cf. Wari', Cavineña, Tariana). One may argue that Tense is still obligatory because the temporal meaning is there, even though the formal marking is not.

Klein (2009, 41) wonders "whether tense and aspect are not completely superfluous in view of what temporal adverbials allow us to do". Languages should satisfactorily be able to compensate the lack of morpho-syntactic Tense marking with the use of adverbials, i.e. it is expected that in Tenseless languages there are more time adverbials than in Tensed ones. This is apparently not true. Bohnemeyer (2009, 114) points out that in Yucatec, a Tenseless language, time can be and indeed is usually expressed without using any time adverbials. But the essential argument, namely that if Tense marking is not *per se* necessary to express temporal relations why some languages (obligatorily) do have it remains unanswered. Bohnemeyer (ibid. 123) points out that Tense marking serves to "disambiguate and to facilitate reference resolution" and that "there is a certain division of labor between pragmatics and the functional category system and a tradeoff between expressed and unexpressed categories", but this yet has to be proven for Tenseless SAILS.

Fusion, i.e. cumulative markers, occurs much more frequently in those languages with maximum or near-maximum Tense marking than in those with less Tense markers; this applies to categories both inside and outside (e.g. person, number, animacy etc.) the realm of TAME.

As demonstrated in section 3.3, deictic spatial markers and possibly also directionals contribute to the expression of temporal relations in Pilagá, which does not have morpho-syntactic Tense marking (cf. also Krasnoukhova 2012, 245). Since spatial reference can be extended to express temporal reference, it would be interesting to see why this connection between the concepts of space and time is made in a speaker's mind and if it works both ways. Concerning the latter point, evidence from cognitive experiments "showed that mental representation of duration and displacement are asymmetrically dependent on one another. Judgments of temporal duration depended on information about spatial extent, but not the other way around" (Casasanto & Boroditsky 2008, 591). This suggests a cognitive hierarchy of space over time, which is linguistically mirrored by the fact that there is no

evidence for a language that has temporal markers expressing spatial relations. This hypothesis certainly needs further analysis.

Only three languages cannot be accounted for by the Tense systems outlined in section 3.4 (Urarina, Trumai, and Mapuche). This supports the idea that grouping SAILs into Tenseless, FUTURE VS. PAST, FUTURE VS. NONFUTURE, PAST VS. NONPAST, and a three-way distinction is in general a valid one and can serve as a basis for further typological studies. It is not claimed that the groups above represent the only way to group languages according to Tense, but for the sample chosen it presents the most natural one since it derives directly from language-particular observations. It should be borne in mind that the data is indeed flattened to achieve a level of comparison and that the individual language Tense systems may not match the groupings completely.

The remainder of this chapter discusses the genealogical and geographical implications and how they relate to previous statements. Section 3.7.3 aims to give a tentative explanation for the heterogeneity of Tense marking, and section 3.9 summarizes the results.

Figure 3.3: Tense categories marked by SAILs

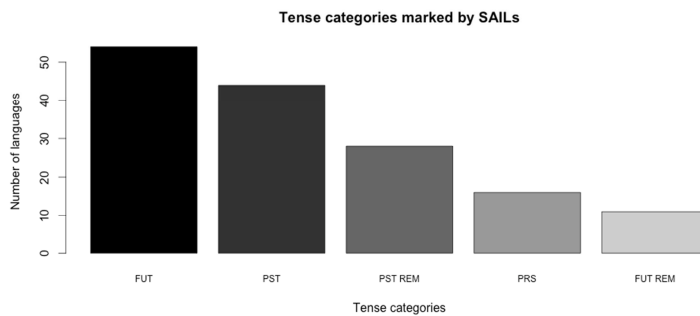


Figure 3.4: Number of SAILs marking number of Tense categories

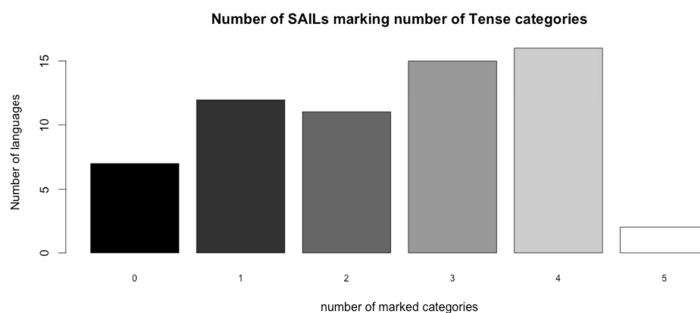


Table 3.7: Overt Tense marking in the SAILS (dark= marked)

	LANGUAGE	FAMILY	FUT	PST	PST REM	PRES	FUT REM	TOTAL
1	Baure	Arawakan						0
2	Tsafiki	Barbacoan						0
3	Mocoví	Guaycuruan						0
4	Pilagá	Guaycuruan						0
5	Nasa Yuwe	Paezan						0
6	Yanam	Yanomaman						0
7	Urarina	Unclassified						0
8	Apurinã	Arawakan						1
9	Chimila	Chibchan						1
10	Bororo	Macro-Géan						1
11	Kaingang	Macro-Géan						1
12	Timbira	Macro-Géan						1
13	Sateré-Mawé	Tupían						1
14	Kamaiurá	Tupían, Guaraní						1
15	Munichi	Unclassified						1
16	Nheengatú	Tupían, Guaraní						1
17	Itonama	Unclassified						1
18	Yurakaré	Unclassified						1
19	Mapuche	Araucanian						1
20	Eméillon	Tupían, Guaraní						2
21	Mosetén	Mosetenan						2
22	Ika	Chibchan, Aruak						2
23	Paresi	Arawakan						2
24	Awa Pit	Barbacoan						2
25	Huallaga Quechua	Quechuan						2
26	Imbabura Quechua	Quechuan						2
27	Cuzco Quechua	Quechuan						2
28	Kanoê	Unclassified						2
29	Yanesha'	Arawakan						2
30	Trumai	Unclassified						2
31	Puinave	Unclassified						3
32	Rikbaktsa	Macro-Géan						3
33	Jarawara	Arawan						3
34	Tehuelche	Chonan						3
35	Aguaruna	Jívaroan						3
36	Wichí (Mataco)	Matacoan						3
37	Cavineña	Tacanan						3
38	Movima	Unclassified						3
39	Karo	Tupían						3
40	Kwaza	Unclassified						3
41	Sabanê	Nambikwaran						3
42	Karitiana	Tupían						3
43	Warao	Unclassified						3
44	Dâw	Nadahup						3
45	Mekens	Tupían						3

46	Tariana	Arawakan						4
47	Panare	Cariban						4
48	Aymara	Aymaran						4
49	Hixkaryana	Cariban						4
50	Tiriyó	Cariban						4
51	Wari'	Chapacuran						4
52	Mamaindê	Nambikwaran						4
53	Matses	Panoan						4
54	Desano	Tucanoan						4
55	Leko	Unclassified						4
56	Shipibo-Konibo	Panoan						4
57	Miraña	Boran						4
58	Hup	Nadahup						4
59	Yaminahua	Panoan						4
60	Cocama-Cocamilla	Tupían, Guaraní						4
61	Tapiete	Tupían, Guaraní						4
62	Cubeo	Tucanoan						5
63	Embera	Chocoan						5
			54	44	28	16	11	

3.8 GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE

There are so far no detailed comparative Tense studies for SA language families, regions, or the whole SA continent that I am aware of.³ In this section I attempt to give a first impression of whether there are distinct Tense patterns geographically, which is complemented by the genealogical study in 3.9. Because it is not always possible to separate geographical and genealogical relationships, there is a certain overlap in these sections.

It was already stated in section 3.6.9 that remoteness degree marking occurs significantly higher in SA than in the rest of the world; but there is no significant frequency of either PAST or FUTURE compared to a global sample: the study of inflectional FUTURE in a global sample of 222 languages by Dahl & Velupillai (2011a) yields 110 marked and 222 languages unmarked for FUTURE. In the present study, two thirds of the languages overtly mark FUTURE which may lead to the conclusion that FUTURE is marked more frequently in SA than elsewhere, but this may be due to a wider definition of 'marker'; Dahl & Velupillai's study only includes affixes and they disregard clitics, particles, and auxiliaries. A closer look reveals that of the 44 languages marking FUTURE in SA, 37 languages utilize affixes (predominantly suffixes) which occur in only slightly more than half of the sample. Thus, although FUTURE is the most frequently marked Tense in the sample, its number in relation to the number of globally marked FUTURE is not significantly higher.

For the PAST Tense, a global sample shows 88 languages that do not mark PAST by inflection and 134 that do mark PAST. In SA, six languages do not have PAST marking, and 15

³ The WALS features 65, 66, 67, and 68 present a most welcome first broad overview, though without detailed discussion.

do (Dahl & Velupillai 2011b). Of the 63 languages in the present sample, 27 mark PAST through affixation and 36 do not mark PAST inflectionally. So whereas in Dahl & Velupillai's (2011b) sample PAST in general and in SA specifically is marked by two thirds of all languages, in the present sample it is marked by less than half. This may be due to the fact that in the present sample zero markers are not taken into consideration, and/ or to a discrepancy in the definitions of PAST and PERFECTIVE.

Table 3.8: FUTURE and PAST inflectional marking in WALS (2011) and Mueller (present/absent (total number of languages analyzed))

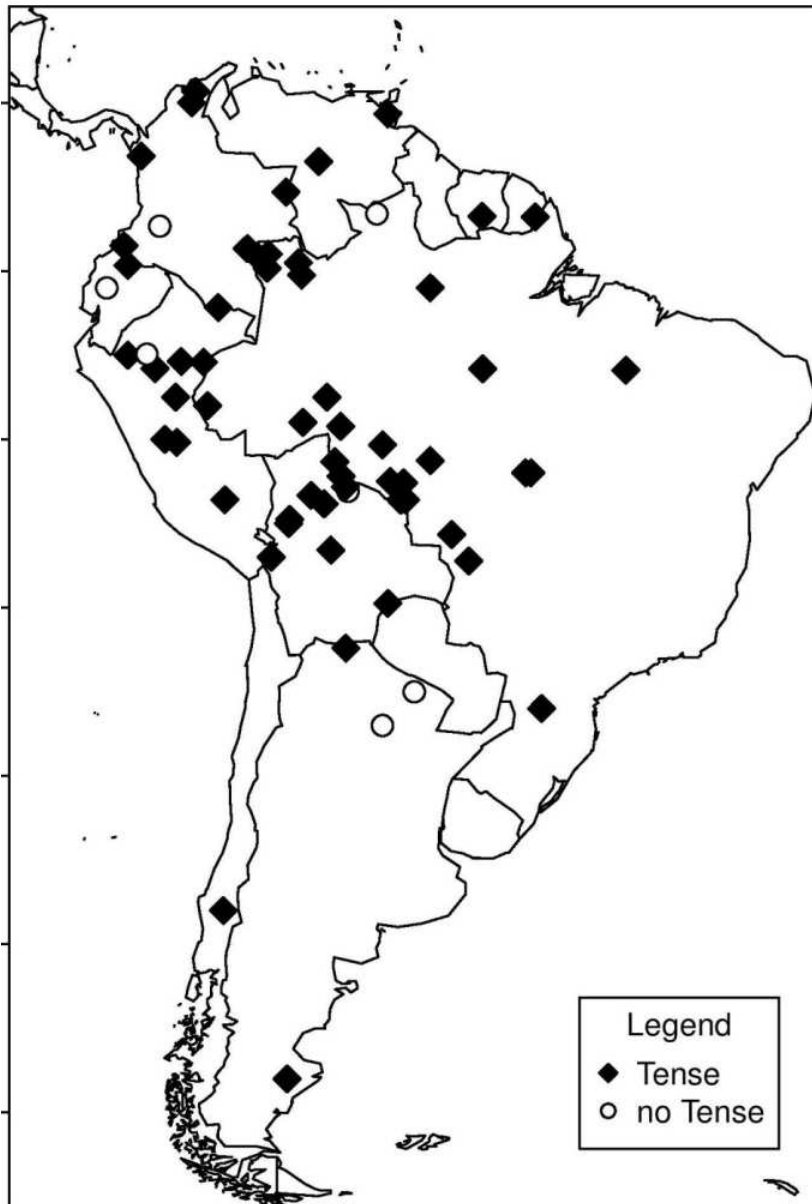
study	FUTURE		PAST	
	SA	global	SA	global
Dahl & Velupillai 2011a, 2011b	10/11 (21)	110/112 (222)	15/6 (21)	134/88 (222)
Mueller (present volume)	37/26 (63)	-	27/36 (63)	-

The geographical distribution of the three Tense categories PRESENT, PAST, and FUTURE is mostly inconclusive. The dispersal of FUTURE in South America is widespread and occurs almost everywhere; the same holds true for PAST and Tense marking in general. Overt PRESENT marking is scattered in the northern half of the continent. Although it is scarce everywhere, it is conspicuously absent along the Andean range.

Very few clear patterns of Tense marking emerge. A possible explanation of this heterogeneity is discussed in 3.10, but first I will turn the focus towards genealogical relationships of Tense marking in the sample. The following paragraphs present an overview and detailed discussion of Tense in Macro-Gêan, Arawakan, Tupían, Cariban, and Quechuan.

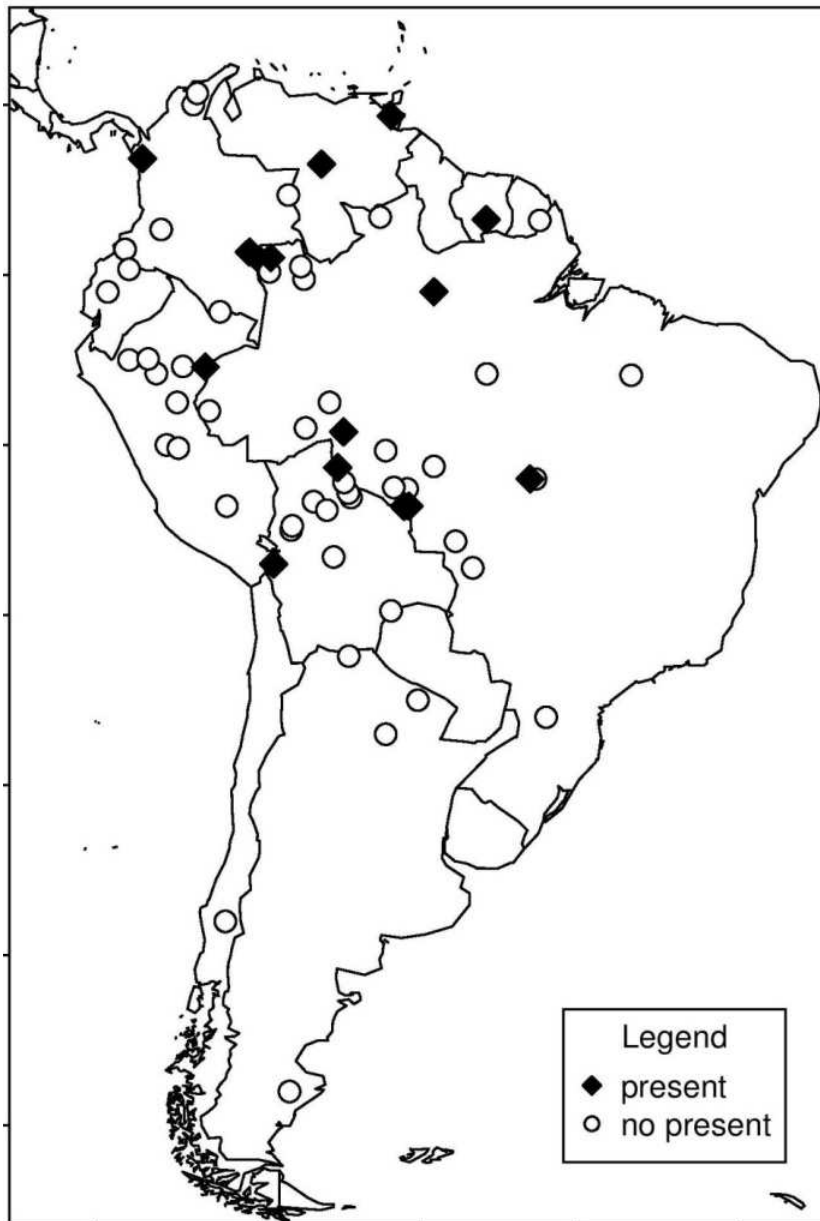
Map 3.2: Distribution of Tense marking in the sample

Tense marking



Map 3.3: Distribution of PRESENT marking in the sample

Present



3.9 GENEALOGICAL DISTRIBUTION IN THE SAMPLE

3.9.1 Introduction

In terms of genealogy, Tense marking within language families is sometimes varied and sometimes homogeneous. The Guaycuruan, Nambikwaran, Nadahup, and Quechuan languages in the sample are relatively homogeneous in terms of which Tense categories they mark. This may be misleading as there are only two languages of each family (three for Quechuan) in the sample, and the variety may be considerably higher once languages are added. It should be borne in mind, that the Tense systems are not *per se* identical, but that sharing Tense systems means a certain amount of shared underlying Tense structures. Quechuan and to a certain degree Tupían and Macro-Gêan show similarities that may be due to common ancestry. Especially Tense marking in Quechuan seems to be very persistent (see section 3.9.5 below). The following sections present Tense marking as occurring in the sample of selected language families: Macro-Gêan (3.9.2), Arawakan (3.9.3), Tupían (3.9.4), Cariban (2.9.5), and Quechuan (3.9.6).

3.9.2 Macro-Gêan

Stassen (1997, 460) already observed for Bororo that “it seems safe to assume that categories of time reference play a subordinate role in the verbal system”, and this principle holds for the other Macro-Gêan languages in the sample as well. Macro-Gêan Tense morphology is poor and usually only FUTURE is marked (see Arawakan for a similar distribution), with the exception of Rikbaktsa, which marks PAST. Two of the four Macro-Gêan languages in the SAILS sample have the same Tense systems: Bororo and Kaingang have FUTURE (marked)/ NONFUTURE (unmarked); Rikbaktsa has as PAST/ NONPAST binary split, and Timbira has a FUTURE/ PAST system. Although Alves (2004) does not give a PAST marker for Timbira, Rodrigues (1999) lists several examples of Timbira that show a marker *tɛ* glossed with ‘ergative PAST’. The same marker is glossed only as ‘PAST’ later on:

(3.40) Timbira (Macro-Gêan, Rodrigues 1999, 181)

(a) *i tɛ pĩ.co j-ũʔkʰər*
 1SG ERG.PST tree.fruit CNT-buy
 ‘I bought fruit.’

(b) *i tɛ amji pĩʌr*
 1SG PST REFL defend
 ‘I defended myself.’

Alves (2004) glosses the same form as ‘ergative’ with no reference to Tense:

(3.41) Timbira (Macro-Gêan; Alves 2004, 108)

ta tɛ kuhi Ø pĩr
 rain ERG fire ABS to.put.out
 ‘The rain put out the fire.’ (orig. ‘a chuva apagou o fogo’)

That *tɛ* is a PAST marker is supported by the FUTURE/ PAST system found in Canela-Krahô which has forms almost identical with Timbira: recent PAST marker *te*, FUTURE marker *ha/ -kra* and a distant PAST marker *pê* (Popjes & Popjes 1989, 179-180).

Table 3.9: Tense in Macro-Gêan

Languages	Tense system	Remoteness degree(s)	Forms
Bororo	FUTURE (marked)/ NONFUTURE (unmarked)	-	FUTURE enclitic (?) <i>modü</i>
Timbira	FUTURE/ PAST	-	FUTURE IRREALIS particle <i>ha</i> , distant PAST <i>pe</i> , PAST <i>tɛ</i> (?)
Kaingang	FUTURE (marked)/ NONFUTURE (unmarked)	-	FUTURE suffix <i>-j/</i> verbal ending in vowel; FUTURE particles <i>kej, ke, jé/</i> verbal root ending in consonant
Rikbaktsa	PAST/ NONPAST binary split	-	NONPAST transitive singular object or 2nd plural: <i>pi-~p-</i> ; NONPAST transitive object 1st or 3rd plural: <i>m-</i>

3.9.3 Arawakan

Tense systems in the Arawakan language family are very considerably varied: Baure is Tenseless, Apurinã and Yanasha' have FUTURE (marked)/ NONFUTURE (unmarked), Tariana has a three-way split, and Paresi has FUTURE/ PAST. According to Aikhenvald (1999, 93-94), many North-Arawakan languages only distinguish between PAST and NONPAST with the exception of Tariana, which obligatorily marks for Evidentiality as a result of contact with Tucanoan. As Tariana is the only North-Arawakan language in the sample, this cannot be confirmed, but Tariana certainly is exceptional in this sample as well.

According to Derbyshire (1986, 529), Brazilian Arawakan languages “do not exhibit any evidence of a tense category in which past, present and future distinctions are found” and Wise (1989, 586) similarly states that “apart from future [...] tense distinctions are relatively unimportant”. It is true that the Arawakan languages in the sample are rather poor in Tense morphology and that if a language marks Tense it is at least FUTURE. Baure is completely Tenseless, and with the exception of Tariana none of them have more than two Tense markers: Paresi has a FUTURE/ PAST distinction with one FUTURE enclitic *=ite/ te* and one PAST enclitic *=ene/ =n* (cf. also *ibid.* 531). Yanasha' and Apurinã both have marked FUTURE/unmarked NONFUTURE Tense systems, although the forms are not the same: Apurinã has a non-immediate FUTURE suffix *-ko*, and Yanasha' enclitic *=cha'* (and allomorphs *-ch/ -V'ch*) plus an alternative near FUTURE auxiliary *o'ch* which may be related to *-cha'*.

(3.42) Apurinã (Arawakan; Facundes 2000, 410)

nhi-nhipoko-ta-ko

1SG-eat-VBLZ-FUT

'I will do eating.'

(3.43) Paresi (Arawakan; Brandão 2011)

- (a) *ha-fidya halia ite*
 3s-clean.up around-PROG FUT
 'He will be cleaning up the field.'
- (b) *zane maidyat-ita=ene*
 go fish-PROG=PST
 'He was fishing.'

Yanesha' has been heavily influenced by Quechuan as evident in numerous loans, the change from an Arawakan four-vowel system to a Quechuan three-vowel system and certain phonological changes (Adelaar with Muysken 2004, 424). This is possibly also observed in the Tense system. Although Yanesha's Tense system is in accordance with the other Arawakan systems in the sample with overtly marking only future, the Yanesha' future marker =*cha*' looks strikingly like the Quechuan future markers. Because the other Arawakan future markers in the sample do not exhibit a similar form it is very likely that Yanesha' acquired its future marker from Quechuan. It is unknown, however, whether it replaced another future marker or was an innovation.

The Arawakan languages in the sample show diversity in Tense system marking but are overall poor in Tense morphology. When Tense is marked, it is at least FUTURE; if a second Tense is marked, it is PAST. Tariana does not fit into this pattern any more due to contact with Tucanoan. The markers are apparently not related to proto-forms, but all have the form of suffixes.

Table 3.10: Tense in Arawakan

Language	Tense system	Remoteness degree(s)	Forms
Baure	Tenseless	-	-
Apurinã	FUTURE (marked)/ NONFUTURE (unmarked)	-	FUTURE suffix - <i>ko</i>
Paresi	FUTURE/ PAST	-	FUTURE enclitic = <i>ite/ te</i> ; PAST enclitic = <i>ene/ ne</i>
Tariana	Three-way split	2 in the PAST	Fused paradigm of PRESENT, PAST and Evidentiality; FUTURE suffixes - <i>mhade</i> and - <i>de</i>
Yanesha'	FUTURE (marked)/ NONFUTURE (unmarked)	1 in the FUTURE	FUTURE enclitic = <i>cha</i> ' (and allomorphs - <i>ch/ -V'ch</i>), immediate auxiliary <i>o'ch</i>

3.9.4 Tupían

Although Tupían has received meticulous attention and is one of the best described language families in South America, there has been no detailed study of the Tense profile of Tupían so far (cf. Everett 2006, 267). The following section gives an overview of Tense occurring in the

Tupían languages in the sample, and may serve as a starting point for an exhaustive Tupían Tense profile.

There are nine Tupían languages in the sample, with five of the Guaraní sub branch (see table 3.11). The most frequently marked Tense category is FUTURE followed by PAST; only Karitiana overtly marks PRESENT (in a cumulative morpheme with PAST). Tense marking ranges from fairly simple (e.g. Kamaiurá with one FUTURE particle) to quite elaborate (e.g. Cocama-Cocamilla with remoteness degrees in PAST and FUTURE). Four languages show remoteness marking. Those languages with a more elaborate Tense system (i.e. a fair number of markers) usually show a bound Tense system, but those with a low number of markers have unbound markers. For example, in Karitiana (six markers) and Tapiete (four markers) all markers are suffixes, whereas in Sateré-Mawé (two markers), Nheengatú (one marker), and Kamaiurá (one marker) all are particles. The exceptions are Karo and Mekens as they both mark Tense by about five particles.

There is a very low degree of correspondence among the markers; only a few Tense markers seem to be related, which will now be discussed. If at some point there have been proto-Tupían Tense forms, they have diverged to a degree that makes it impossible to reconstruct them. However, there is slight evidence for a Tupían PAST suffix.

In Karitiana and Mekens a suffix *-t* occurs. In Mekens, this (non-obligatory) marker expresses PAST in transitive and intransitive main clauses (it changes to *-r* when followed by a vowel) (Galucio 2001, 90–91). The situation in Karitiana is more complicated: *-t* expresses NONFUTURE when preceded by a vowel in copula constructions, speech-act-participant constructions, valence constructions with intransitive verbs, and is obligatory (its allomorph *-n/ -in* occurs in verb focus constructions (Everett 2006, 265). There is no similar PAST or NONFUTURE marker in the other Tupían languages in the sample. However, Rodrigues (1999, 116) cites Gabas Jr. (1994) in two examples with an ominous Aspectual suffix *-t*. When looking at these examples, it is clear that the *-t* suffix rather expresses a NONFUTURE than Aspectual value:

(3.44) Karo (Tupían; Rodrigues 1999, 116, citing Gabas Jr.)

- (a) *na?to to-wirap ?o-t*
 tapir 3SG.C-food eat-ASPECT
 ‘The tapir ate its own food.’
- (b) *na?to a?-wirap ?o-t*
 tapir 3SG.NC-food eat-ASPECT
 ‘The tapir is eating its [something else’s] food.’

However, Gabas Jr. (1999, 73) does not mention this suffix in the Tense, but rather in the Modality section, treating it as INDICATIVE. Interestingly, the allomorph of INDICATIVE *-t* is *-n* (following nasals), which again resembles the allomorph of the Karitiana NONFUTURE allomorph *-n* (in verb focus constructions). Some support for the hypothesis that the Karo suffix *-t* is a Tense marker comes from Moore (1994):

- (3.45) Karo (Tupían; Moore 1994, 159)
iyít w-e-t a-ma-wiy-a
 squeeze 1SG-AUX-PST(?) 3SG-CAUS-go.out-PTCP
 ‘I squeezed it out (referring to a foot worm).’

Other Tupían languages that have a PAST suffix *-t* are Wayoro (with allomorph *-n*) and Ayuru:

- (3.46) Wayoro (Tupían; Galucio & Nogueira 2011, 23, 24)
 (a) *õn kiβi pikarẽη-k-a-t*
 1SG stick crooked-VBLZ-THEM-PST
 ‘I bent the stick.’
 (b) *ngwajkip iβoj-tikwa-p nõ-ø-ã-n te-ndaip mẽ*
 man fish(N)-fish(V)-NOM give-VBLZ-THEM-PST 3-son PSPT
 ‘The man gave a fishing net to his son.’

- (3.47) Ayuru (Tupían; Moore 1994, 153)
o-kip agwa gora-t
 1SG-brother cará look.for-PST
 ‘My brother looked for cará.’

Although probably the reflex of a Tupían NONFUTURE/ PAST proto-form the suffix *-t* in Mekens, Wayoro, Karo, Ayuru, and Karitiana needs further study.

The FUTURE markers of the Tupían languages apparently are not related, but Jensen (1998, 536) claims that for the Guaraní branch of Tupían the verb **potár* ‘want’ is commonly used for FUTURE (and DESIDERATIVE). This is evident in the Tapiete immediate FUTURE suffix *-pota* and it may also occur in the Emérillon FUTURE marker *-tal* (Rose 2003, 426), but not in the other Guaraní languages in the sample (Mueller 2013).

The Tense systems of the Tupían languages in the sample are FUTURE marking, either in combination with PAST or NONFUTURE, but besides that show little correspondence. Besides a possible NONFUTURE/ PAST suffix **-t* in Mekens, Karitiana, and Karo, and traces of **potal* in FUTURE markers in Tapiete and Emérillon, the Tense markers of the Tupían languages in the sample seem to be unrelated.

In contrast, Jensen (1999, 128) notes Guaraní languages are “noted for a high degree of lexical and morphological similarity among its member languages in spite of their extensive geographical separation”. This is not true for Tense as Rodrigues (1999, 118) has stated that “the languages vary a good deal in how they mark categories of aspect, tense and mood”. Tense marking in Tupían is thus largely heterogeneous with the exception of having mostly FUTURE/ NONFUTURE distinctions.⁴

⁴ This also is true for Awetí (only Tense marker is the FUTURE particle *tut*) (Drude 2008, 71).

Table 3.11: Tense in Tupían

Language	Family	Tense system	Remoteness degree(s)	Forms
Sateré-Mawé	Tupían	FUTURE (marked)/ NONFUTURE (unmarked)	-	FUTURE particles <i>aru, wuat</i>
Karitiana	Tupían	FUTURE/ NONFUTURE binary split	-	cumulative suffixes PRESENT and PAST (NONFUTURE) <i>-t, -n, -in</i> , FUTURE suffixes <i>-i, -j, -ndaki</i>
Karo	Tupían	FUTURE/ PAST	2 in the PAST, 1 in the FUTURE	2 FUTURE particles <i>yat, iga</i> ; 1 FUTURE auxiliary <i>kap</i> ; 2 PAST particles <i>co, kán</i> (recent and remote) (PAST suffix <i>-t?</i>)
Mekens	Tupían	FUTURE/ PAST	1 in the FUTURE	FUTURE particles <i>pek, pa, paat, kot</i> ; cumulative FUTURE particle with irrealis (<i>p)egat</i> ; PAST suffix <i>-t/-r</i>
Emérillon	Tupían, Guaraní	FUTURE/ PAST	-	FUTURE suffix <i>-tal</i> , immediate PAST suffix <i>-o?u</i>
Kamaiurá	Tupían, Guaraní	FUTURE (marked)/ NONFUTURE (unmarked)	-	FUTURE particle <i>korin</i>
Cocama- Cocamilla	Tupían, Guaraní	FUTURE/ PAST	3 in the PAST, 2 in the FUTURE	immediate PAST enclitic <i>=uy</i> , mediate PAST enclitic <i>=ikud</i> , remote PAST enclitic <i>=tsuri</i> , immediate CERTAIN FUTURE enclitic <i>=utsu</i> , mediate/ remote FUTURE enclitic <i>=d</i>
Tapiete	Tupían, Guaraní	FUTURE/ PAST	2 in the PAST, 2 in the FUTURE	immediate FUTURE suffix <i>-pota/ -po</i> , distant FUTURE suffix <i>-kwi</i> , immediate PAST suffix <i>-ye ~ -e</i> , distant PAST suffix <i>-kwe</i>
Nheengatú	Tupían, Guaraní	FUTURE (marked)/ NONFUTURE (unmarked)	-	Future particle <i>kurí</i> ⁵

3.9.5 Cariban

In this sample, Cariban languages are characterized by their high degree of fusion between Tense, Aspect, Modality, and number. For example, in Tiriyo the indicative verb receives affixes of two different sets, either factual or nonfactual. The factual set is divided into IMPERFECTIVE and PERFECTIVE which is further subdivided into collective/ non-collective and each has different forms for PRESENT, PAST, FUTURE. Additionally, there is a DUBITATIVE vs. CERTAINTY distinction in the PRESENT and FUTURE. The nonfactual set is divided into collective and non-collective, as is the distinct set of forms for IMPERATIVE (Meira 1999, 281). The

⁵ Possibly related to *kury* 'action or state at point of change' in Guajajára (Jensen 1998, 554), and *kuri* 'recent past, attested' in Kaiwá (ibid.).

situation in Hixkaryana is similar, although without the factual/ non-factual distinction, NONPAST instead of distinct PRESENT and FUTURE, and three remoteness degrees instead of one (Derbyshire 1985, 196). All three Cariban languages have an Aspectual distinction in most PAST forms between COMPLETIVE/ CONTINUATIVE OR IMPERFECTIVE/ PERFECTIVE and a Modal distinction in NONPAST/ PRESENT and FUTURE forms (cf. Derbyshire 1999, 38 and Gildea 1992, 27–28).

According to Derbyshire (1999, 38), Cariban languages have remoteness degrees in the PAST but not in the FUTURE, and this is confirmed for the present sample. Hixkaryana, Panare, and Tiriyo all have remoteness degrees only in the PAST.

In contrast to many other language families in SA relatively much is known about Tense in Modern Cariban and Proto-Cariban, not least because of Gildea (1998). Gildea (ibid. 98) reconstructs several Proto-Cariban TAM forms, reflexes of which can also be seen in this data, and argues that despite some variation they descend from a proto-system:

”Having compared the syntax, verbal personal prefixes, TAM suffixes, and number suffixes/ particles that constitute the Set1 system in 19 modern languages, despite some language-specific variation we can see that the overall morpho-syntactic patterns are so consistent that they can only be explained as having descended from a single protosystem” (ibid. 101).

As can be seen in table 3.12, Hixkaryana is probably still most faithful to the proto-forms and Tiriyo has changed the most (the data for the three languages found in Gildea differ to a minor degree from that presented here). There is no distinction any more between recent and remote PAST, but whereas Hixkaryana and Panare have cumulative NONPAST forms, Tiriyo has separate markers for FUTURE and PRESENT. An innovation seems to be the prefix/ suffix combination *te-* *-se* in Tiriyo that marks remote PAST, and which is mostly used in narratives of myths and legends. This is not an Evidential, because it marks that the event is far removed from the point of speech rather than indicating how the information was obtained. A component which all three languages share is an Aspectual distinction in the PAST. Because Gildea’s study dates back almost 20 years, it would be interesting to see whether the reconstructed proto-forms stand up when more recent data are added.

Table 3.12: Tense distinctions in Cariban (present sample) and Proto-Cariban (Gildea 1998)

			Tense forms in the present sample ⁶		
			Hixkaryana	Panare	Tiriyó
Gildea' (1998) Tense proto- forms	recent PAST	*-wi/-i/ -no	-no, txowi, -yako, -yatxoko, -yaknano, -yatxkenano	-ya, -i, -e	-(ja)kě(ne), -ne, (te- se)
	distant PAST	*-ne	-ye, txowni, -yakoni, -yatxkoni	-ya(ke)	
	NONPAST	*-ya-ce	-yaha, -yatxhe	-n, -ñe	-(ja)-e, -(ja)-ñe, -ta-e, -ta-(ne), -(ja)-kě(mi)
	NONPAST uncertain	*-ya-no	-yano, -yatxowi	?	

Table 3.13: Tense systems in Cariban

Language	Tense system	Remoteness degree(s)
Hixkaryana	PAST/ NONPAST binary split	3 in the PAST
Panare	PAST/ NONPAST binary split	3 in the PAST
Tiriyó	Three-way split	1 in the PAST

3.9.6 Quechuan

The three Quechuan languages in the sample are very consistent in their Tense marking. Huallaga, Imbabura, and Cuzco Quechua all have FUTURE/ PAST Tense systems and similar forms with only some variation in the FUTURE paradigm. There is no overt marking for PRESENT. A verb with person marking and without FUTURE or PAST marking is interpreted as PRESENT. The PAST Tense suffixes *-ra* (Huallaga), *-rka* (Imbabura), and *-rka* (Cuzco) (see table 3.14) obviously derive from *-rqa (cf. Adelaar with Muysken 2004, 223, cf. Weber 1989, 99) (for this form as PERFECTIVE marker see 4.7.2). Cuzco Quechua additionally has a suffix *-sqa* that also has past time reference. Faller (2002, 30) dismisses earlier analyses of the two Cuzco PAST markers coding an Evidential distinction, but rather identifies *-rka* as normal PAST marker and *-sqa* as “perfect of evidentiality” although she states that the latter marker is not Evidential. Although it occurs largely in narratives with reportative values it is not obligatory in all narratives and occurs also without reportative, but can have only perfect⁷ meaning (as well as mirative and even NON-FIRSTHAND meaning) (ibid. 30-32).

⁶ For ease of reading the different suffixes for number, transitivity, and Aspectual distinctions are listed but not further specified. NONPAST in Tiriyó includes markers for PAST and for FUTURE; the Tiriyó forms under recent PAST are not specific for any remoteness.

⁷ Faller's and Weber's term “perfect” is conform to ANTERIOR as used in this thesis.

(3.48) Cuzco Quechua (Quechuan; Faller 2002, 30, 31)

- (a) *para-sha-sqa* not experienced by speaker
rain-PROG-sqa
'It was raining.'
- (b) *Marya-qa hamu-sqa* PERFECTIVE
Marya-TOP come-sqa
'Marya has come.'
- (c) *pisi-lla-ña yaqa tawa saku-cha* story-telling
little-LIM-DISC almost four sack-DIM
hin-lla-ña muhu ka-sqa
like-LIM-DISC seed be-sqa
'There was only a little, almost only four sacks of seeds.'
- (d) *kay-pi-má ka-sha-sqa Marya-qa.* mirative (surprise)
this-LOC-má be-PROG-sqa Marya-TOP
'Marya is here.'

Huallaga Quechua has a similar marker *-sha/ -shka* which Weber (1989, 196) classifies as “perfect” Tense, and which gradually takes over as a general PAST marker, whereas *-ra* is used as a narrative marker. Weber (ibid. 106, footnote 7) further argues that *-shka* and *-sha* stem from an original participle+auxiliary construction (*-shqa-ka*) which signaled perfect and collapsed into *-sha/ -shka*. So whereas in Huallaga Quechua *-ra* is in the process of shifting towards a marker with Evidential value and the original perfect marker takes over PAST functions, in Cuzco Quechua the original PAST marker *-rka* retains its past time reference and the perfect marker *-sqa* has Evidential values. This is interesting, because it demonstrates that two related languages have the same semantic shift although with different source markers. No such shift is observed in Imbabura Quechua, where the suffix *-shka* remains a perfect and *-ra* a PAST marker (Cole 1982, 144–145), but, according to Hintz (2007), the former perfect (here: ANTERIOR) suffix *-shqa* in South Conchucos Quechua also changed into a recent PAST marker.

(3.49) Huallaga Quechua (Quechuan; Weber 1989, 100)

- (a) *aywa-ra-n* narrative PAST
go-PST-3
'He went.'
- (b) *aywa-sha* “everyday” PAST
go-3PFTV
'He went.'

South Conchucos Quechua exhibits a remoteness distinction in the PAST, which is not observed for the other three Quechuan languages (according to the sources). The former

perfect marker *-shqa* now codes recent PAST and the original PAST marker *-rqa* depicts a more remote PAST. There are actually two recent PAST markers which occur in complementary distribution according to person: *-shqa* and *-rqu* (Hintz 2007, 39).

FUTURE in the Quechuan languages does not have a single marker like PAST, but uses a distinct person paradigm with almost identical forms in all three languages. The FUTURE paradigms are defective, i.e. there is no difference between PRESENT and FUTURE person marking in the second person. In Huallaga Quechua, a FUTURE form is often accompanied by the suffix *-paq* which means ‘at some future time’, but this does not seem to be obligatory (Weber 1989, 100).

(3.50) Huallaga Quechua (Quechuan; Weber 1989, 100)

- (a) *aywa-nki*
go-2SG
‘You go.’ and ‘You will go.’
- (b) *aywa-shaq-kuna-paq*
go-1FUT-PL-FUT
‘We (excl.) will go.’

The relation between Cuzco Quechua first person FUTURE *-sqa* and the PERFECTIVE marker *-sqa* is unknown. It is likely that the FUTURE stems from proto-Quechuan form, because it is so similar to the first person FUTURE in Huallaga and Imbabura Quechua, but it remains to be seen whether the PERFECTIVE marker is somehow derived from this or an independent development.

It may seem illogical to analyze FUTURE as overtly coded, but not PRESENT as both do not have distinct suffixes like the PAST but rather a person paradigm. Why take the FUTURE person paradigm as cumulative markers but the PRESENT one as unmarked? The answer is that because the same person markers in the PRESENT are used in the PAST and thus do not carry PRESENT Tense value as such. Additionally, the person markers of the FUTURE paradigm always refer to future time reference.

The Quechuan languages are quite homogeneous in Tense marking. Neither marks PRESENT, but does mark PAST and FUTURE, and that with almost the same sets of markers. They undoubtedly stem from proto-Quechuan forms with very little variation.

Table 3.14: Tense marking in Quechuan (Huallaga, Imbabura, Cuzco)

Language	PRESENT	PAST	FUTURE					
			1 st sg/pl		2 nd sg/pl		3 rd sg/pl	
Huallaga	-	<i>-ra/ (-sha/ -shka)</i>	<i>-shaq</i>	<i>-shun</i>	-	-	<i>-nqa</i>	<i>-nqa</i>
			<i>(+ -paq)</i>					
Imbabura	-	<i>-rka</i>	<i>-sha</i>	<i>-shun</i>	-	-	<i>-nqa</i>	<i>-nqa</i>
Cuzco	-	<i>-rka/ (-sqa)</i>	<i>-sqa</i>	<i>-sqa</i>	-	-	<i>-nqa</i>	<i>-nqa</i>

3.10 STABILITY OF TENSE

A major feature of Tense marking in the SAILS is the great heterogeneity of the Tense markers in the sample. There are only a few instances where Tense markers resemble each other within language families. One possible explanation is a high rate of change and the related property of instability, which has been proposed on the basis of an analysis of the WALS features by Wichmann & Kamholz (2008) and Wichmann & Holman (2009). The basic assumption behind grading stability of typological features draws on the apparent incongruity of change of different features.⁸ Stability itself refers to the susceptibility of a feature to change and change refers to the transformation of phonological material of a feature. If there is a difference in two markers for the same feature of two related languages then change must have happened. Regarding the present data, establishing Tense proto-forms is very difficult for most families as the markers diverge very much (except for Quechuan) which points towards a high rate of change. This in turn points toward low stability. The following section introduces the measuring of stability of typological features and how this relates to the present study. In a study of WALS data the relative stability of PAST and FUTURE were measured by Wichmann & Holman (2009) and the results are now compared to the present data.

According to Wichmann & Holman (2009, 45), Tenses in the WALS sample vary according to their diachronic stability: the feature “past tense” is very stable, but “future tense” unstable. This is based on a tested stability metric that is based on the distinction between related and unrelated languages. The authors take stability of a feature “defined as the probability that a given language remains unchanged with respect to the features during 1000 years” (ibid. 12). For the two features above this implies that in the WALS sample PAST is very unlikely to change within 1000 years and FUTURE likely to change. A previous test by Wichmann & Kamholz (2008) revealed that PAST Tense ranks in the lower half of stability of (WALS) features, and FUTURE Tense very low. There seemingly is a discrepancy in the ranking of PAST in both studies, but because Wichmann & Holman consider the metric of the first study performing better than the second one I will not discuss the second one here.

The ranking of stability relates to the probability of change, but the study shows two more refined results that are interesting for the present discussion. It is not only interesting to know to which degree features are prone to change, but also which changes are involved. Feature changes can be internal, i.e. due to typological language internal developments, or external, i.e. due to contact-induced change. The latter, called “diffusion” by Wichmann & Holman, occurs in the most stable features and in turn internal change does not affect the most stable features as much as relatively less stable ones: “stable features are more resistant to internal change but no more resistant to diffusion than are unstable features.[...] the most stable features change almost entirely by diffusion, with negligible

⁸ See Wichmann & Holman (2009, 5) for a historical account of stability in typology.

internal change” (Wichmann & Holman 2009, 42, 38). Thus, if this is true, then PAST and FUTURE in this sample should both be affected by diffusion, but FUTURE more than PAST by internal change. Acting on the assumption that the stability study of Wichmann & Holman (2009) is valid, we can conclude two working hypotheses for this sample:

1. PAST should show a relatively high stability, i.e. a low degree of change within language families, and when there is change, it should be conditioned relatively more frequently by diffusion than internal change.
2. FUTURE should show a high degree of change, i.e. FUTURE marking is relatively heterogeneous in languages of the same family, and this change is due to both diffusion and internal change.

The danger with statements about proto-forms is that there is no possibility to prove them, but this is irrelevant to the argument of stability. When related languages mark the same categories and the markers do not share significant phonological material, they are either internal/ external innovations or have changed to a degree that the original cannot be retraced and are thus not stable. In the following paragraphs, language families with more than one member in the sample are tested according to the two hypotheses above.

The marking of the category PAST in Cariban is stable, but the forms appear to be only moderately stable. According to the data gathered by Derbyshire (1999, 38, table 2.7), there are re-occurring PAST forms throughout Apalaí, Bakairí, Carijona, Dekwana, Hixkaryana, Panare, Tiriyó, Wai Wai, and Wayana, although not to a satisfying degree. For example, Apalaí, Hixkaryana and Dekwana mark PAST or immediate PAST with variations of *-no*, but Bakairí with *-da* and Panare with *-yah*. On the other hand, NONPAST is marked with variations of *-ya* in almost all of these languages, which suggests that NONPAST is more stable in Cariban than PAST.

The second hypothesis is supported by the data in Macro-Gêan and Tupían: Macro-Gêan does not mark PAST, but all languages in the sample do mark FUTURE. The forms suggest that a common ancestor form is unlikely: Bororo *modü*, Rikbaktsa *m-/p-*, Kaingang *-j, kej/ ke/ ké*; Timbira *ha*. Although a single proto-form may have existed, the markers have changed to a very high degree. Tupían FUTURE forms are similarly diverse: Karo *yat, iga, kap*; Mekens *pek, pa, paat, kot*, Sateré-Mawé *aru, wuat*; Emérillon *-tal*; Kamaiurá *korin*; Cocama-Cocamilla *=utsu, =á*, Tapiete *-pota/ -po, kwi*; Nheengatú *kuri*. Besides the above mentioned source **potar* ‘want’ for FUTURE in some Guaraní languages (Timbira and Emérillon) the Tupían FUTURE markers are not related.

The data from Tupían and Macro-Gêan supports the claim that FUTURE is very unstable, and from Cariban that in comparison PAST is moderately stable. An exception to both hypotheses is Tense in Quechuan which is very stable; the PAST as well as FUTURE forms are almost identical in all three languages in the sample (see section 3.9.6). To test whether the heterogeneous markers are due to diffusion or internal change more research needs to be done on the individual markers which is beyond the scope of this study. It can be concluded

that the high heterogeneity in Tense markers may be explained by its relative stability or lack thereof as proposed by Wichmann & Holman (2009).

But the heterogeneity may also be due to the fact that TAME markers exist in a continuum, i.e. markers are prone to shift from one category to another. In the realm of Tense, this is especially the case for FUTURE and Epistemic Modality, or PAST and PERFECTIVE. For stability of Aspect see section 4.11, for Modality 5.13, for Evidentiality 6.6, and for a final discussion see section 7.5.

3.11 SUMMARY

This chapter investigated the typological, geographical, and genealogical patterns of Tense marking in the sample. The languages were analyzed according to FUTURE, PRESENT, PAST, and remoteness degrees in the PAST and FUTURE in 63 SAILS. Tense was defined as the grammaticalized expression of time and each category was further defined as the relationship between the points of event and speech. It was demonstrated that morpho-syntactic Tense marking in the SAILS sample can be divided into Tensed and Tenseless languages. Tensed languages fall into the classes of two-way splits (FUTURE/ NONFUTURE simple split, FUTURE/ NONFUTURE binary split, PAST/ NONPAST binary split), three-way splits (FUTURE, PRESENT, PAST), and a FUTURE/ PAST division. The latter occurs most frequently with 22 members, followed by FUTURE/ NONFUTURE simple split (14), three-way split (nine), PAST/ NONPAST binary split (six), and FUTURE/ NONFUTURE binary split (three).

Most of the languages in the sample have morpho-syntactic Tense marking, only seven are Tenseless (and three are borderline cases). Of the possibility to mark up to five Tense distinctions (PRESENT, FUTURE, PAST remote PAST, remote FUTURE), 12 languages mark one, 11 mark two, 15 mark three, 16 mark four, and two mark five Tenses. The most frequently marked category is FUTURE, which is marked by 54 languages, followed by PAST (44), remoteness in the PAST (28), PRESENT (16), and remoteness in the FUTURE (11). This allows for a hierarchy of Tense marking in the sample: FUTURE > PAST > PRESENT. There is no significant frequency of marking of PAST or FUTURE compared to a global sample in Dahl & Velupillai (2011b); however, the marking of remoteness degrees in SA is considerably higher than in the rest of the world. No significant correlation between Tense systems and remoteness degrees or any geographical pattern could be found.

A major characteristic of morpho-syntactic Tense marking in the SAILS is its heterogeneity of forms which may be explained by the relatively low stability of FUTURE and PAST marking as calculated in a previous study. Thus, all Tense categories are apparently highly susceptible to change, but in general FUTURE more so than PAST. Tense marking in all categories is varied in all language families (with more than two members) in the sample with the noticeable exception of Quechuan. The Quechuan languages are quite homogeneous in their Tense marking: neither marks PRESENT, but all mark PAST and FUTURE with cognates. In contrast, Arawakan, Macro-Gêan and Tupían languages, although exhibiting similar systems, are widely inconsistent in their Tense makers. Cariban Tense

marking is in general is more homogeneous than in Arawakan, Tupían and Macro-Gêan, but less than in Quechuan.

4. ASPECT

4.1 INTRODUCTION

In contrast to Tense, which situates events on a timeline, Aspect is concerned with the internal structure of the event. Thus, Aspect expresses the ‘how’ of an event rather than the ‘when’: “[a]spect is not concerned with relating the time of the situation to any other point in time, but rather with the internal temporal constituency of the one situation” (Comrie 1976, 5). Klein (see section 3.1) lists Aspect as one means to express temporal distinctions in languages, and while it is true that Tense and Aspect are related (cf. Dahl 1985, Comrie 1976, Bhat 1999 etc.) this chapter focuses on the viewpoint of actions rather than their temporal situation. This is not always separable as e.g. cumulative markers of Tense and Aspect show (e.g. PAST in Tiriyo is divided into PERFECTIVE and IMPERFECTIVE), so Tense will be discussed when relevant as well.

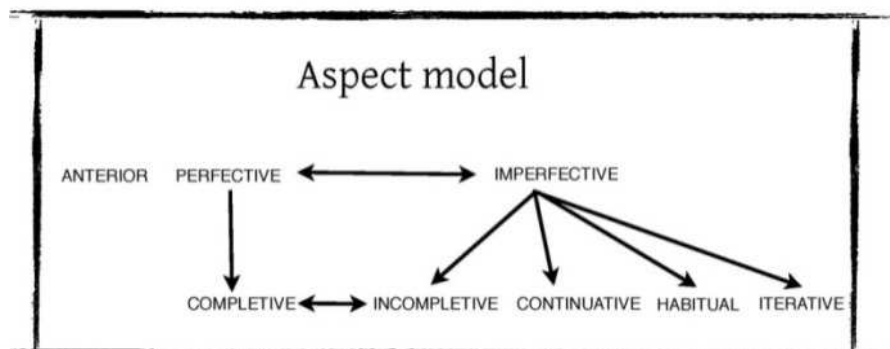
Under Aspect I understand the internal characteristics of actions as “viewpoint” in contrast to situation types, also called “Aktionsart”. The latter is not part of the present investigation, which will focus on viewpoint. Thus, the term Aspect as used here refers to viewpoint only. Aktionsart refers to the inherent Aspectual properties of verbs. Givón (2001, 287-288) classifies Aktionsart verbs as short duration (also called punctual) verbs, accomplishment verbs, activity (also called dynamic) verbs, and stative verbs. These properties are relatively stable, i.e. usually a verb within a language does not change its inherent status and is not dependent on viewpoint. Viewpoint on the other hand varies from speaker to speaker, depending on his view, and can thus in theory be applied to all Aktionsart verbs. Nevertheless, languages apparently have certain restrictions as to which viewpoint can be used with which class of Aktionsart, which complicates the analysis of viewpoint Aspect. Just as Aspect is not completely separable from Tense, viewpoint cannot be completely kept separate from Aktionsart, which means that in this chapter Aktionsart plays a role as well, although the focus is on viewpoint Aspect.

Different from the Tense categories outlined in chapter 3, Aspectual categories are not that easily delimited from each other. PRESENT, PAST, and FUTURE intuitively are different concepts and can be separated from each other by the formulas based on Reichenbach (1947), but in order to arrive at definitions to keep Aspectual markers apart, it is necessary to refer to the prototype/ canonical theory (see chapter 2). As a starting point, a model will be developed, based on previous studies, which will be used as a working hypothesis. It is based on the study of Bybee et al. (1994), who in turn based their study on the works of Comrie (1976), Dahl (1985), and Givón’s continued work about Tense and Aspect. Whereas Comrie’s and Givón’s work is more of a theoretical nature (but nevertheless based on genuine data), Bybee et al.’s and Dahl’s studies represent global quantitative analyses. This has the advantage that the definitions are based on two major quantitative studies, and that they reflect actual data against which the present data can be mapped. Another advantage is that the WALS chapter about PERFECTIVE/ IMPERFECTIVE is also based on Dahl (1985) and Bybee et al. (1994), enabling a comparison with even more languages.

The Aspectual categories chosen for the questionnaire are PERFECTIVE, IMPERFECTIVE, ANTERIOR, HABITUAL, CONTINUATIVE, ITERATIVE, COMPLETIVE, and INCOMPLETIVE. These categories have been chosen because they represent the major Aspectuals as found in typological literature so far. HABITUAL and ANTERIOR have sometimes been categorized as Tense in previous studies, but are nevertheless put under Aspect here, because their prototypical meanings, outlined in the relevant sections below, relate more to Aspect than to Tense.

Comrie (1976, 25) assumes a binary opposition of PERFECTIVE and IMPERFECTIVE and breaks IMPERFECTIVE further down into the oppositions HABITUAL and CONTINUOUS, and CONTINUOUS into PROGRESSIVE and NONPROGRESSIVE. Comrie (1976, 24f.) likewise observes that “while many languages do have a single category to express imperfectivity, there are other languages where imperfectivity is subdivided into a number of distinct categories, and yet others where there is some category that corresponds to part only of the meaning of imperfectivity”. Bybee et al. (1994, 139) “adopted this view of imperfective and have coded as imperfective only those forms with uses in both habitual and continuous contexts”. The latter approach is followed here with the following modifications: IMPERFECTIVE is indeed a supercategory, comprising HABITUAL, CONTINUATIVE, ITERATIVE, and INCOMPLETIVE. For a form to be IMPERFECTIVE, it needs to have at least the meaning of two of the four possible subcategories. PERFECTIVE is in opposition to IMPERFECTIVE, COMPLETIVE is a subcategory of PERFECTIVE, and COMPLETIVE and INCOMPLETIVE are binary oppositions. ANTERIOR is a category separate from PERFECTIVE and IMPERFECTIVE. This working model does not claim to be universal or exhaustive, but it serves as a starting point to classify and analyze the Aspectuals of the languages in the sample.

Figure 4.1: Aspect model



Each section below gives a list of characteristics that have to be met in order for a marker to be classified as encoding the respective Aspectual values. The Aspect model entails that most of the Aspectuals studied here share meanings to a certain extent, so it is necessary to include negative characteristics, such as ‘does not include a habit or custom’ for CONTINUATIVE and ITERATIVE to be separated from HABITUAL. Bybee et al. (1994, 138) regard

negative lists as inappropriate, “since the diachronic development of grams shows that they have real semantic content”. I agree, and that is why the definitions of the different Aspects include negative and positive characteristics.

The present analysis closely relies on descriptive grammars and articles, but sometimes deviates from the analysis given in those sources, e.g. in Tsafiki: according to Dickinson (2002, 233), Tsafiki expresses PERFECTIVE and IMPERFECTIVE with suffixes that form relative clauses; in main clauses there is no (IM)PERFECTIVE marking. Because this study focuses on Aspectuals in main clauses, Tsafiki does not have (IM)PERFECTIVE marking in this study.

The SAILS show a variety of Aspect markers. All SAILS mark at least one of the following: PERFECTIVE, IMPERFECTIVE, PERFECT, HABITUAL, CONTINUATIVE, ITERATIVE, COMPLETIVE, OR INCOMPLETIVE. The languages differ in the degree to which category is marked, how it is marked formally, and which combinations of marking occur within a language; nevertheless, it is possible to make a few generalizations which will be discussed in 4.8.

This chapter is structured as follows: Section 4.2 focuses on the PERFECTIVE/IMPERFECTIVE distinction. The sections after that examine in more detail the marking, meaning and form of HABITUAL (4.3), CONTINUATIVE (4.4), ITERATIVE (4.5), COMPLETIVE/ INCOMPLETIVE (4.6), and ANTERIOR (4.7). Sections 4.8 to 4.10 aim to consolidate the findings and discuss the implications for Aspect typology (4.8), geographical (4.9), and genealogical (4.10) relationships. The implications of the results with regard to typological stability of Aspect are discussed in 4.11. Section 4.12 summarizes the results. Where relevant, the points of contact between Tense, Aspect, Aktionsart, and Modality will be mentioned.

4.2 PERFECTIVE/IMPERFECTIVE

4.2.1 Definitions

PERFECTIVE and IMPERFECTIVE are two opposing viewpoints. Whereas PERFECTIVE refers to an action viewed a single whole, IMPERFECTIVE allows for internal structuring. “The term ‘perfective’ contrasts with ‘imperfective’, and denotes a situation viewed in its entirety, without regard to the internal temporal constituency” (Comrie 1976, 12) whereas IMPERFECTIVE is “viewing a situation from within” (ibid. 24). Dahl (1985, 78) specifies PERFECTIVE to “denote a single event, seen as an unanalyzed whole, with a well-defined result or end-state. More often than not, the event will be punctual, or at least, it will be seen as a single transition from one state to its opposite, the duration of which can be disregarded”. An event marked with PERFECTIVE is usually interpreted as being temporally bounded, as in Apurinã:

- (4.1) Apurinã (Arawakan; Facundes 2000, 517)
- | | | | | |
|--------------|-----------------|--------------|--------------|------------------|
| <i>oposo</i> | <i>uwa-kata</i> | <i>su-pe</i> | <i>o-txa</i> | <i>hãtako-ro</i> |
| after | 3SG.M-ASO | GO-PFV | 3F-AUX | youth-F |
- ‘Then the young woman went with the tapir.’

PERFECTIVE can include the meaning of COMPLETIVE, i.e. that the focus is on the termination of the action (see section 4.6).

IMPERFECTIVE refers to an action that is not viewed in its entirety but with subparts. These subparts can have HABITUAL, INCOMPLETIVE, ITERATIVE, and/ or CONTINUATIVE meaning. An IMPERFECTIVE marker can have all of these but must at least have two meanings. For example, in Awa Pit the suffix *-tu* can mark CONTINUATIVE or HABITUAL:

(4.2) Awa Pit (Barbacoan; Curnow 1997, 225, 227)

- (a) *profesora=ta titizh-tu-s*
 teacher=ACC wait.for-IPFV-LCT
 'I am waiting for the teacher.'
- (b) *mes=ayzpha Ricaurte=ta puz-tu-s*
 month=every Ricaurte=in go.out-IPFV-LCT
 'Every month I go out to Ricaurte.'

4.2.2 Distribution in the sample

In the sample, 11 of the languages mark PERFECTIVE, 10 IMPERFECTIVE, and 13 mark both (regardless of which other Aspects are marked) (see table 4.1). Languages in the sample that morpho-syntactically mark PERFECTIVE, but not IMPERFECTIVE (regardless of other Aspect marking) are: Imbabura Quechua, Cuzco Quechua, Warao, Trumai, Miraña, Desano, Sateré-Mawé, Apurinã, Baure, Dâw, and Hup. Languages in the sample that morpho-syntactically mark IMPERFECTIVE, but not PERFECTIVE (regardless of other Aspect marking) are: Embera, Kanoê, Wichí, Aguaruna, Awa Pit, Puinave, Paresi, Nasa Yuwe, and Cavineña. Languages which mark both are: Chimila, Yanam, Munich, Kaingang, Hixkaryana, Panare, Nheengatú, Yurakaré, Yaminahua, Shipibo-Konibo, Tiriyo, Leko, and Huallaga Quechua.

The form of marking ranges from suffixes over to particles and auxiliaries, with a preference for suffixes. Only Yurakaré is prefixing. When both categories are marked, the markers usually appear in the same clause position, i.e. PERFECTIVE and IMPERFECTIVE are both marked by suffixes, or by particles, or else, and mutually exclusive (regardless of what other means that language has to express (IM)PERFECTIVE)¹. For example, in Munich PERFECTIVE is marked by the suffix *-me* and IMPERFECTIVE by suffix *-mu*:

(4.3) Munich (unclassified; Gibson 1996, 53, 54)

- (a) *tu?name?nu manse?sana*
 arrive-PFV-1SG church-LOC
 'I have already arrived at the church.' (orig. 'Ya he llegado a la iglesia.')
- (b) *u?samura*
 dark-IPFV-3SG
 'It is already dark.' (orig. 'Ya está oscuro.')

¹ See chapter 3.7 for a discussion of imbalance in formal Tense marking.

Yurakaré is the only language in the sample prefixing (IM)PERFECTIVE (or any other Aspect). Additionally, the PERFECTIVE prefix *i-* is connected to verbal number. It often co-occurs with distribution marking (*-uma* or reduplication of final verb syllable), but still marks events that are temporally bounded. In (4.4b), the action of searching is done repeatedly in different places, every single instance being bounded.

(4.4) Yurakaré (unclassified; Van Gijn 2006, 188, 189; Van Gijn p.c.)

- (a) *a-bobo-∅*
 IPFV-hit;kill-3
 ‘He is fighting.’
- (b) *i-bëbë-bë-∅*
 PFV-search~DISTR-3
 ‘He searched everywhere.’

A number of languages mark PERFECTIVE/ IMPERFECTIVE in certain Tenses, but not in others. For example, Tiriyo has a PERFECTIVE/ IMPERFECTIVE opposition in all its marked Tenses, i.e. FUTURE, PRESENT, PAST (although PRESENT PERFECTIVE is unmarked), but in Hixkaryana PERFECTIVE/ IMPERFECTIVE marking is restricted to recent and distant PAST.

(4.5) Tiriyo (Cariban; Meira 1999, 152, 300, 181, 304, 307)

- (a) *mëe_pë n-ëturu-ja-n*
 3AN.PROX_about 3S_A-talk-PRS.IPFV-DUB
 ‘S/he is talking about this one.’
- (b) *t-ëpëi ipahka-∅*
 3-seat break-PRS.PFV
 ‘S/he has just broken his/her seat.’
- (c) *ëremina-në w-eta-ne*
 sing-GINF 1A-hear-PST.PFV
 ‘I heard singing.’
- (d) *ji-npo ki-ri-ta-e*
 I-on.back 1+2A-do-FUT.IPF-CERT
 ‘I will put you on my back.’
- (e) *suurinam_po w-ei-ne, wi-tën-ne serë_pëë*
 Surinam-LOC 1S_A-COP-PST.PFV 1S_A-go-PST.PFV 3INA.PROX_ABL
 ‘I was in Surinam (then, long ago), I went there from here.’

(4.6) Hixkaryana (Cariban; Derbyshire 1979, 136)

- (a) *r-otaha-yako*
 3SBJ.1OBJ-hit-REC.PST.PFV.NCOL
 ‘He hit me.’

- (b) *oy-owakrye-yatxkoni*
 3SBJ.2OBJ-make happy-DIST.PST.IPFV.COL
 ‘They made you happy.’
- (c) *ni-emen-yakoni*
 3SBJ-steal-DIST.PST.IPFV.NCOL
 ‘He used to steal.’
- (d) *r-otaha-yaha*
 3SBJ.1OBJ-hit-NPST.NCOL
 ‘He will hit me.’

It is not unusual for a language in the sample to have a marked PERFECTIVE, an unmarked IMPERFECTIVE and also a marker for a category that is in general considered unbounded, i.e. belonging to the IMPERFECTIVE sphere. Apurinã, Panare, Hup, Desano, Trumai, and Warao all have an unmarked IMPERFECTIVE, but do mark at least HABITUAL. It is therefore more precise to say that a general IMPERFECTIVE in those languages may be unmarked, however there does exist marking for certain IMPERFECTIVE inherent categories.

In Apurinã, PERFECTIVE is marked with the suffix *-pe*. It expresses temporal boundedness and various secondary meanings of PERFECTIVE: COMPLETIVE, RESULTATIVE, and INCHOATIVE (change of state). In addition, *-pe* can co-occur with the FUTURE marker. Facundes (2000, 515ff.) also claims that *-pe* marks anterior (PERFECT). IMPERFECTIVE, on the other hand, is unmarked, and it seems that in order to fill this gap Apurinã employs HABITUAL and INCOMPLETIVE suffixes (but not CONTINUATIVE).

The WALS map for PERFECTIVE/ IMPERFECTIVE Aspect (feature 65) presents 21 languages for the South American continent. Languages coded either have grammatical marking of PERFECTIVE/ IMPERFECTIVE or not. In contrast to the present study, periphrastic constructions are included (Dahl & Velupillai 2011a). Table 4.2 presents the languages coded for PERFECTIVE/ IMPERFECTIVE in WALS and the given values, and which values were coded for this study. From 21 languages in WALS, nine are also present in the SAILS sample. Two languages have conflicting data: Wichí and Tucano.

The ratio of PERFECTIVE/ IMPERFECTIVE marking vs. no marking is 5:16 for the WALS data in SA, and 33:30 for my data. Considering the total number of languages coded for this feature in WALS (222), almost half of them (101) mark PERFECTIVE/ IMPERFECTIVE morphologically and/ or periphrastically. The ratio of the world’s PERFECTIVE/ IMPERFECTIVE marking is therefore about the same for South America, but because my data does not include periphrastic

constructions, it is expected that the ratio is even higher in the SAILS if one adds periphrastic constructions.

Table 4.1: PERFECTIVE/ IMPERFECTIVE marking in the SAILS

	PERFECTIVE only	IMPERFECTIVE only	Both
Languages	Imbabura Quechua, Cuzco Quechua, Warao, Trumai, Miraña, Desano, Sateré-Mawé, Apurinã, Baure, Dâw, Hup	Embera, Kanoë, Wichí, Aguaruna, Awa Pit, Puinave, Paresi, Nasa Yuwe, Cavineña	Chimila, Yanam, Munichí, Kaingang, Hixkaryana, Panare, Nheengatú, Yurakaré, Yaminahua, Shipibo-Konibo, Tiriyo, Huallaga Quechua, Leko

Table 4.2: PERFECTIVE/ IMPERFECTIVE marking vs. no marking in SAILS

Language	PFV/IPFV in WALS ²	PFV/IPFV in Mueller
Warao	No	Yes
Sanuma	No	-
Tucano	No	No (Cubeo) Yes (Desano)
Awa Pit	Yes	Yes
Imbabura Quechua	No	No
Jivaro	No	-
Yagua	No	-
Barasano	No	-
Hixkaryana	Yes	Yes
Pirahã	Yes	-
Apurinã	Yes	Yes
Chácobo	Yes	-
Araona	No	-
Wari'	No	No
Canela -Krahô	No	-
Cochabamba Quechua	No	-
Wichí	No	Yes
Abipón	No	-
Guaraní	No	-
Mapuche	No	No
Cocama	No	-

4.3 HABITUAL

4.3.1 Definition

HABITUAL markers express an action that takes place regularly over a period of time, i.e. a custom or habit. HABITUAL should not be confused with ITERATIVE which expresses repetition, but not custom (Comrie 1976, 27). The action in question may be customarily repeated/or be in a state over a certain amount of time: "they describe a situation which is characteristic of an extended period of time, so extended in fact that the situation referred to is viewed

² From Dahl & Velupillai 2011a.

not as an incidental property of the moment but, precisely, as a characteristic of a whole period” (ibid. 27–28), e.g. in Pilagá:

- (4.7) Pilagá (Guaycuruan; Vidal 2001, 269)
*qo-ya-sona-**pega***
 IPS.SBJ-3-stick-HAB
 ‘They used to stick it by the fire.’ (e.g. meat or fish, to be roasted)

4.3.2 Distribution in the sample

Of the 63 languages in the sample, 32 morpho-syntactically mark HABITUAL:

Jarawara, Mapuche, Tapiete, Bororo, Mekens, Cubeo, Pilagá, Yanasha’, Urarina, Imbabura Quechua, Puinave, Mamaindê, Warao, Trumai, Kwaza, Miraña, Embera, Kaingang, Nasa Yuwe, Desano, Sateré-Mawé, Panare, Apurinã, Nheengatú, Yurakaré, Yaminahua, Shipibo-Konibo, Tiriyo’, Tariana, Dâw, and Hup.

Formal marking ranges from a simple suffix (e.g. Mapuche), cumulative suffixes (e.g. Cubeo) over clitics (e.g. Tariana) to particles (e.g. Mekens). Many of the SAILs without HABITUAL marking have an IMPERFECTIVE marker that expresses HABITUAL, and even those with HABITUAL markers can have additional IMPERFECTIVE markers that express HABITUAL. It is certainly not uncommon to find more than one means to express HABITUAL morpho-syntactically in one language in the sample.

Habituality can be expressed by an IMPERFECTIVE marker, as in e.g. Cavineña and Awa Pit. In Cavineña, the IMPERFECTIVE marker can express (besides others) HABITUAL, but exclusively with past time reference; Awa Pit does not restrict HABITUAL to the past:

- (4.8) Cavineña (Tacanan; Guillaume 2008, 173)
*yawa=ju=shana ekana ani-**ya**.*
 ground=LOC=PITY 3PL sit-IPFV
 ‘They would sit (directly) on the ground, the poor women.’

- (4.9) Awa Pit (Barbacoan; Curnow 1997, 227)
*mes=ayzhpá Ricaurte=ta puz-**tu-s***
 month=every Ricaurte=in go:out-IPVF-LOCUTOR.PERSON.MARKER
 ‘Every month I go out to Ricaurte.’

It has been noted before that HABITUAL forms can also be used in generic sentences (cf. Dahl 1985, 99). This is confirmed for the SAILs as well, for example in Tariana, Cubeo, and Jarawara. The Tariana enclitic *-hyuna* ‘habitual prescribed’ marks generic statements (only occurs in impersonal constructions) (Aikhenvald 2003, 326). In Cubeo, the PRESENT HABITUAL

forms also mark for general truths (Morse & Maxwell 1999, 41). Indeed, many of the HABITUAL markers in the sample extend their meaning to a generic one.

(4.10) Tariana (Arawakan; Aikhenvald 2003, 327)

diha pusaru-ne hyukade-mha
 ART sloth-FOC.SBJ not.appear-PRS.NVIS
di-pitana-mia phema-hyuna-mha
 3SG.NF-name-ONLY IPS+hear-HAB-PRS.NVIS
 ‘The sloth does not appear, one only hears his name.’

(4.11) Cubeo (Tucanoan; Morse & Maxwell 1999, 44)

aru ā-lj-Abē kũ-wA-RE
 and eat-STV-PRS.HAB/NREC.PST.M.SG worm-PL-OBJ
 ‘And he (a fish) habitually eats worms.’

4.4 CONTINUATIVE

4.4.1 Definition

The following characteristics have to be met by a marker to be CONTINUATIVE:

1. The verbal action is ongoing, usually, but not necessarily, during the point of reference.
2. No custom or habit is involved.
3. In contrast to HABITUAL, when happening for an extended period of time, a CONTINUATIVE action must take place during the whole time, whereas HABITUAL actions during that time would occur with interruptions.

These points have to be met regardless of the Aktionsart of the verb, although they may be restricted to a certain Aspect class. For example, in Apurinã the CONTINUATIVE suffix *-nanu* cannot occur on descriptive verbs, but still is considered a CONTINUATIVE marker. Jarawara is an example for a language where the CONTINUATIVE marker *-ne* (sometimes realized as zero morpheme) can occur on both dynamic and stative verbs.

(4.12) Apurinã (Arawakan; Facundes 2000, 525)

- (a) *n-umaka-nanu-ta*
 1SG-sleep-CNT-VBLZ
 ‘I’m sleeping.’
- (b) **here-nanu-ta*
 be.pretty-CNT-VBLZ
 ‘He’s being handsome.’

(4.13) Jarawara (Arawan; Dixon 2004, 187, 188)

- (a) *moto waka na*
 motor(M) be.broken AUX+CNT.M
 ‘The motor is broken.’
- (b) *tera noki o-ne o-ke*
 2NSG.OBJ wait.for 1SG.SBJ-CNT.F 1SG-DECL.F
 ‘I am waiting for all of you.’

4.4.2 Distribution in the sample

Of the 63 languages in the sample, 35 show morpho-syntactic CONTINUATIVE marking:

Mocoví, Karitiana, Jarawara, Mapuche, Tsafiki, Wari’, Itonama, Tehuelche, Sabanê, Ika, Bororo, Mekens, Cubeo, Pilagá, Yanasha’, Urarina, Imbabura Quechua, Puinave, Rikbaktsa, Emérillon, Cocama-Cocamilla, Mosetén, Paresi, Timbira, Munichich, Cuzco Quechua, Nasa Yuwe, Desano, Sateré-Mawé, Panare, Apurinã, Aymara, Baure, Dâw, and Leko.

Formal marking ranges from suffixes (e.g. Imbabura Quechua), cumulative suffixes (e.g. Cubeo), over clitics (e.g. Mocoví) to particles (e.g. Timbira) and auxiliaries (e.g. Dâw). Most of the SAILs do not overtly mark Tense distinctions with CONTINUATIVE meaning. This may be due to the fact that an ongoing action is not inherently impossible in either FUTURE, PRESENT, or PAST. In languages which do not overtly or obligatorily mark Tense it is simply impossible to establish a restriction, but it may also be due to e.g. constituent structure: in Karitiana, the CONTINUATIVE suffix *-(a)ri* is not restricted to Tense, because due to constituent order Tense clitics and *-(a)ri* rarely co-occur (Vallejos Yopán 2010, 342). The suffix is also a “real” CONTINUATIVE in Comrie’s sense as it occurs on all types of verbs; mental processes, states, and both telic and atelic verbs.

In Wari’, CONTINUATIVE marking is restricted to PAST Tense. Although it is possible to express PRESENT PROGRESSIVE/ CONTINUATIVE by e.g. the unmarked verb, the only CONTINUATIVE particle in Wari’ is restricted to PAST:

(4.14) Wari’ (Chapacuran; Everett & Kern 1997, 316)

- quep nana-in xirim ’ira*
 do 3PL.REA.PST/PRES-3N house PST.CNT
 ‘They were making the house.’

Pilagá makes a distinction of marking HABITUAL on lexically durative verbs and states. The first is marked by the PROGRESSIVE suffixes *-tak/ -ta* (allomorphs) and *-tape* (third person PROGRESSIVE) (Vidal 2001, 260). When a stative verb is required to become ongoing, the durative suffix *-tapiñi/ -tapiyi* is added to the verb. Vidal (ibid. 263) argues that the durative suffix has a different meaning from PROGRESSIVE, “since many activity verbs can alternatively

take *-tak* and *-tapiñi*, and the difference in meaning is based on whether the event is conceptualized as effectively going on (i.e., with *-tak*), or as having a certain duration, without reference to a time frame (i.e., with *-tapiñi*)”.

(4.15) Pilagá (Guaycuruan; Vidal 2001, 264)

(a) *na-losos-tapiñi*

3-run-DUR

‘He runs/ ran and runs/ran.’

(b) *ña-losos-tak*

3-run-PROG

‘He is running.’

4.5 ITERATIVE

4.5.1 Definition

The term ITERATIVE as used in this study comprises several meanings that together form the ITERATIVE prototype. Unlike e.g. PRESENT, there is not a single semantic characteristic, but it is possible to identify a most important one, i.e. the meaning of an action that is repeated. An action that is not repeated disqualifies for ITERATIVE, but actions that are repeated are not automatically ITERATIVE. DISTRIBUTIVE and HABITUAL also mark actions that are repeated, but these can take place over an extended period of time with relatively long intervals between them. An ITERATIVE marker applies to actions that are repeated instantly. When a marker only marks repetitions happening over a longer time period with extended intervals this points toward HABITUAL or DISTRIBUTIVE meaning. For example, the suffix *-dyi* in Mosetén only refers to repetitions with relatively long periods of time between them and therefore does not match the definition of ITERATIVE, although in every other way it behaves like it:

(4.16) Mosetén (Mosetenan; Sakel 2004, 270)

kaph-kaph-dyij.

clap-REDUP-ITE.M.SBJ

‘He claps (constantly).’

The following characteristics must be present for a marker to have ITERATIVE meaning:

1. The verbal action is repeated (regardless of how many times).
2. The same subject(s) carries out the action.
3. There is no custom involved.
4. There is no change in meaning of the verbal action.
5. The repetitions happen immediately after each other (event-internal).

Points one to five constitute a prototypical ITERATIVE. Because this definition is very specific and detailed, it is expected that there are less markers exclusively marking ITERATIVE and

comparably more markers that have ITERATIVE as possible meaning, but also extensions: I will now discuss each of the points in more detail.

Point one is the most important and most obvious characteristic of an ITERATIVE marker and it also relates to point five. Repetitions can take place only once or several times; the exact number does not matter as long as they follow one another immediately (i.e. without intervening actions). Several SAILS feature markers that are usually translated into English as ‘again’; these can be pitfalls. In Mapuche, the suffix *-tu* apparently marking a single repetition, but always refers to an action that restores a former state, or the return of a referent to a former state/ location. This suffix is not ITERATIVE, because it implies a return. It is not the same action that is done again, as e.g. in (4.17a) the setting free is not repeated, and in (4.17b) the returning is not repeated:

(4.17) Mapuche (Araucanian; Smeets 2008, 254, 255)

- (a) *nel-üm-tu-fi-n*
become.loose-CAUS-RE-OBJ-IND.1SG
‘I set him free again.’, ‘I restored his freedom to him.’
- (b) *amu-pe-tu-la-y*
go-PROX-RE-NEG-IND-3
‘He probably went back.’

Sometimes, an action that looks repeated is just a chain of uninterrupted actions. In (4.18), the children walked for an extended period of time, but the walking cannot be broken down to repetitions of doing the same walking. They did not walk the same path over and over again (which in any case would imply that they also returned, and this meaning is excluded from ITERATIVE, see above).

(4.18) Baure (Arawakan; Danielsen 2007, 229)

<i>no=kač-po-wo</i>	<i>to</i>	<i>powor</i>	<i>ahi-nev</i>	<i>kač</i>
3PL=go-PFV.REFL-COP	ART	poor	child-PL	go
<i>no=yono-po</i>	<i>no=yono-po</i>	<i>no=yono-po</i>		
3PL=walk-PFV.REFL	3PL=walk-PFV.REFL	3PL=walk-PFV.REFL		

‘The poor children went away, the[y] walked and walked and walked.’

Point two: The repeated action must be carried out by the same subject or subjects. This is not necessarily the case with e.g. the suffix *-ua* in Matses, which marks an action that can be repeated both by the same or s different subject. In (4.21h) below a different woman than that who did the first strain repeats the straining. However, as there are examples where *-ua* also refers to actions repeated by the same subject and fulfills all of the other characteristics above, this is merely a case of vagueness and *-ua* is an ITERATIVE marker.

Point three: This excludes any action that is carried out by custom and therefore is prototypically HABITUAL. For many examples it is not easy to derive whether the action is

done customarily or not, because context is the best way to establish the real meaning and examples are mostly given out of context. However, the presence of e.g. adverbials such as ‘usually, customarily’ (not in the translation, but the original), or the presence of additional HABITUAL markers in the same clause help.

Point four: ITERATIVITY does not include a change in meaning. Some verbs, when receiving a marker for repetition, change their meaning, e.g. in Cocama- Cocamilla *itika* means ‘throw’ and together with the iterative suffix (*itika-ka*) means ‘separate, get divorced’, instead of ‘throw again/repeatedly’ (Vallejos Yopán 2010, 359). Nevertheless, *-ka* is still coded as ITERATIVE because it also has the ‘normal’ function (see below).

Another meaning deriving from points one, two, and five, is that the repetition takes place at roughly the same location. This is usually a given, because one person can only move so far away before instantly repeating an action. However, because it is possible for certain actions to be repeated while moving (hitting somebody while in a plane, where a considerable distance is covered during the action), it is not a characteristic. It is helpful to keep in mind, though, to distinguish ITERATIVE from DISTRIBUTIVE. DISTRIBUTIVE means that the repetition is carried out by several subjects and/ or on several objects at different locations and over time, and this is not the case with ITERATIVE, where the action is carried out by the same subject(s) and immediately following each other. The Yurakaré suffix *-uma*, although expressing plural actions and therefore repetitions as in (4.20a), also marks one action done by several referents (4.19b):

(4.19) Yurakaré (unclassified; Van Gijn 2006, 190)

- (a) *a-ushpě-wma-∅*
INCP-bathe-DISTR-3
‘He is bathing all the time.’
- (b) *otto-∅=w=ja* *i-bali-wma-∅=w*
go.out-3=PL=SS VPL-go.PL-DISTR-3=PL
‘After they came out (of the water) they each went.’

The most difficult part of analysis is the relationship of ITERATIVE meanings with HABITUAL, CONTINUATIVE, and DISTRIBUTIVE. Comrie (1976, 27) points out that HABITUAL does involve a certain kind of repetition of an action, but that there are two important facts separating ITERATIVE and HABITUAL: HABITUAL includes a) a custom or habit that b) is characteristic for an extended period of time, which both are not true for ITERATIVE. In (4.20), the different instances of beating take place immediately after each other and possibly even simultaneously, forming an action that is conceived as a single entity although it can continue for a while. The beating is also not done habitually, i.e. not repeated after days, months, years, with intervals where no beating takes place.

(4.20) Hup (Nadahup; Epps 2008, 580)

d'ũç tih tətəd-d'óʔ-óy=mah
 timbó 3PL beat(REDUP)-take-DYNM=REP
 'They beat *timbó* (rapidly).'

In cases where the descriptions are too vague and examples too scarce to successfully distinguish between HABITUAL, ITERATIVE, and DISTRIBUTIVE, the marker in question is taken as vague, i.e. marking all or a subset of them. Nasa Yuwe serves as an example. Jung (1979, 65) mentions the effect of reduplication of the last syllable of a verb: in verbs whose action indicates a single action, this means that action is repeated. Unfortunately, no examples in full clauses are given and I must take the author by her word and code reduplication as ITERATIVE.

Rubino (2005, 11) points out that there are two possible types of reduplication: "full vs. partial. Full reduplication is the repetition of an entire word, word stem (root with one or more affixes), or root [...]. Partial reduplication may come in a variety of forms, from simple consonant gemination or vowel lengthening to a nearly complete copy of a base". The questionnaire does not make a distinction between these types, e.g. as long as the phonological material of a verb is either partially or fully reduplicated, and this reduplication marks the repetition of the meaning of the reduplicated verb, it counts as ITERATIVE. In order to be specific when referring to ITERATIVE the term 'reduplication' applies to the form (phonological material) and 'repetition' to the meaning of a verb.

4.5.2 Distribution in the sample

Of the 63 languages in the sample, 32 morpho-syntactically mark ITERATIVE:

Karo, Tsafiki, Wari', Itonama, Tehuelche, Sabanê, Tapiete, Kamaiurá, Aguaruna, Matses, Bororo, Mekens, Cubeo, Rikbaktsa, Emérillon, Cocama-Cocamilla, Mosestén, Mamaindê, Warao, Trumai, Kwaza, Nasa Yuwe, Aymara, Nheengatú, Yurakaré, Yaminahua, Baure, Hup, Tariana, Huallaga Quechua, Cavineña, and Leko.

Of these, 15 mark ITERATIVE only by reduplication of the verb, eight by affixes and/ or particles, and eight have both forms of marking. The following discussion presents examples for each type and illustrates several individual cases of ITERATIVE marking.

A case study of ITERATIVE marking in Matses is now presented to illustrate how the considerations above have been applied to the data. Matses was chosen, because it has several forms that are candidates for ITERATIVE, it demonstrates the difficulties of distinguishing between Aspectual categories, and because the source (Fleck 2003) provides ample data.

Matses has two different morpho-syntactic ways of marking a possible ITERATIVE: suffixation (suffixes *-ded*, *-ban*, *-ua*) and reduplication, and within these distinguishes as much as three meanings that all have in common that they mark repetition. The three

suffixes are all glossed ITERATIVE in (4.21a-h) for the reader to better identify them, although they may be glossed differently throughout the thesis.

In Matses, reduplication of the verbal root (together with certain suffixes) refers to an action that is repeated immediately or over an extended period of time and has almost CONTINUATIVE or HABITUAL appearance (4.21a). Indeed, it is not possible to clearly distinguish between ITERATIVE, HABITUAL, and CONTINUATIVE here; because reduplication often points toward ITERATIVE, it is possible that this meaning was once the origin and that CONTINUATIVE and HABITUAL developed from this. This is certainly an interesting point of overlap between the semantic fields of Aspectuals (in Tapiete, partial verb root reduplication also marks an ongoing action (González 2005, 160)). Interestingly, reduplication of a verb does not necessarily mean repetition of an action; it can also mean that the action is de-intensified, as in (4.21b). This is interesting because one would expect reduplication to mean the exact opposite, i.e. intensifying of the action in question.

Considering (4.21a) and (4.21b), reduplication does fulfill points one to four in the list of characteristics of ITERATIVE above. (4.21a) may be interpreted both as a habit of picking fruits, or as repeatedly picking fruits on one occasion, therefore being ambiguous to point three. In (4.21c), however, the reduplication of *posh* (*poshca* ‘make hole’) points toward immediate repetition, though a HABITUAL meaning seems to be there (note that the verb in the main clause is marked with the HABITUAL suffix *-quid*). Thus, reduplication in Matses fulfills all points of the list but is ambiguous to point three. Fleck (2003, 454) points out that the exact meaning of reduplication varies according to verb transitivity and semantics, the presence of certain directional suffixes and context. It seems that reduplication in Matses is vague according to HABITUAL, ITERATIVE, and DISTRIBUTIVE, but because it includes all of the points in the list above, it is treated as ITERATIVE marker here.

The suffixes suffer from similar vagueness. The suffixes *-ban* and *-ded* fulfill points one, two, and four, but are vague to being customary. The suffix *-ua* is vague according to the immediate or non-immediate repetition. The suffix *-ded* must co-occur with a reduplicated root or the suffix *-ne*, which means distributive, HABITUAL, or plural. The suffix *-ban* can, but usually does not, co-occur with the suffix *-ne*, and can co-occur with reduplication of the verb. The suffixes *-ded* and *-ban* differ in their ITERATIVE meaning: *-ded* implies that there are relatively long intervals of time between the repeated actions, whereas *-ban* refers to actions that are repeated instantly, usually on multiple objects. When it co-occurs with reduplication of the verb, it also means that there are multiple subjects (ibid. 359f.). The suffix *-ua* (homophonous with the verbalizing suffix) means that the repetition can be carried out by a different subject, though that does not have to be the case. In (4.21g), the subject is the same, but in (4.21h) *-ua* marks that the straining is done by a different woman. It is compatible with the adverb *padpide* ‘again’ (ibid. 363f.).

To conclude, none of the three suffixes or the reduplication are unambiguously ITERATIVE. It is possible to exclude the suffix *-ded*, because it does not fulfill point five (immediate repetition), but reduplication and *-ban* are vague according to HABITUAL status,

and *-ua* is vague according to immediate repetition and same/ different subject. Thus, reduplication, *-ban*, and *-ua* are coded as marking ITERATIVE (beside others), and *-ded* is not.

Table 4.3: Matses markers of repetition

Definition/ form	reduplication	<i>-ban</i>	<i>-ded</i>	<i>-ua</i>
Verbal action is repeated	Yes	Yes	Yes	Yes
Same subject carries out the action	Yes	Yes	Yes	Yes/no
No custom	Vague	Vague	Vague	Yes
No change in meaning	Yes	Yes	Yes	Yes
Repetitions happen immediately	Yes/no	Yes	No	Vague
ITERATIVE	Yes	Yes	No	Yes

(4.21) Matses (Panoan; Fleck 2003, 452, 453, 360, 364, 363)

- (a) *chiuish bacuë chedo **cuëstan** cuës-tan-e-c.*
 fig fruit etc. REDUP gather-GO-NPST-IND
 ‘They keep on going to pick fig fruits and similar fruits.’
- (b) **ush** *ush-o-sh.*
 REDUP (hastily) sleep-PST-3
 ‘He slept only a short time.’
- (c) *dëmush usun-quid ad-en dëbiate*
 nose.whisker insert-HAB like.that.MNR:TR nose
*shëcuë-n **posh** poshca-shun.*
 hole-LOC REDUP make.hole-after:SBJ
 ‘They insert the nose whiskers, after repeatedly piecing many holes in their noses like that.’
- (d) *cuesban-n opa **pe-ded** pe-ded-e-c.*
 bat-ERG dog REDUP-ITE bite-ITE-NPST-IND
 ‘The bat(s) keeps on biting the dog(s) every night (the same dog or dogs).’
- (e) *cuesban-n opa **pe-ban** pe-ban-e-c.*
 bat-ERG dog REDUP=ITE bite-ITE-NPST-IND
 ‘Many bats bite many dogs, biting one and then the next.’
- (f) *chu-ua-ua-o-sh.*
 hot-VBLZ-ITE-PST-3
 ‘He heated it up again.’
- (g) *ashumbic pisid padpide-en bed-ua-quid.*
 then woven.mat again-MNR:TR grab-ITE-HAB
 ‘Then, they pick up the mat again.’

- (h) *ashumbic sica-shun utsi-n-tsen ada min secte*
 then strain-after:SBJ other-ERG-next DUB 2GEN strainer
ic-tséc-e-c que-quin bed-uan-shun
 be-DIM-NPST-IND say-while:SBJ grab-come-after:SBJ
aid-n secte-n sica-ua-quad.
 that.one-ERG strainer-INST strain-ITE-HAB
 ‘After that, after she strains, another [woman] comes and asks, “Do you have a strainer?” and then takes it and then she also strains with the strainer.’

The study of Matses above has shown that a language can have multiple markers for ITERATIVE that these interact with other ASPECTUAL values. Other SAILS vary across a range of having no ITERATIVE marking at all and ranging from very simple ways of marking repeated actions to the complexity of Matses.

Reduplication is the most common way to mark ITERATIVE in the sample, either occurring on its own or in addition to another ITERATIVE marker. For example, Cocama-Cocamilla marks ITERATIVE by both reduplication and suffixation. The suffix *-ka* can have the following meanings: repetition over a period of time, with intervals or happening immediately, happening only once or many times, in different locations, and by/ on many subjects/ objects simultaneously (Vallejos Yopán 2010, 382f.). It can also derive verbs from nouns (4.21c). Apart from suffixation, Cocama-Cocamilla marks ITERATIVE by reduplication of either the first two syllables or only the second syllable of the verb; the difference in form is not mirrored in meaning. Both *-ka* and reduplication can, and in some cases must, co-occur (ibid. 369).

(4.22) Cocama-Cocamilla (Tupían-Guaraní; Vallejos Yopán, 2010, 354)

- (a) *raepetsui ta era tikita-ka=ura ikia-ka-tika*
 after 1SG.M be.good tie-ITE=3M.OBJ this=LOC-CER2
 ‘Then I tie it very well (tie multiple times), up to here.’
- (b) *ra yupuni yauki urkuru umi-umi-ka*
 3SG.M start make basket see-see-ITE
 ‘She starts to make the basket looking and looking at the base of paucar (Russet-backed oropendola)’s house.’
- (c) *y=ipu-ta-ka*
 3SG.F=sound-CAU-ITE
 ‘He makes noise/fuss.’ (‘He makes consecutive sounds that became noise.’)

In Aguaruna, it seems to be the interaction of partial reduplication and the suffix *-kawa/ -kua* that marks ITERATIVE. The suffix cannot occur on its own, but must be accompanied by reduplication. Only one example in Overall (2007, 226, ex. 76) shows reduplication of the verb without the suffix.

- (4.23) Aguaruna (Jivaroan; Overall 2007, 383)

buu buuta-kawā wi-u
 REDUP CRY+IPFV-ITE+3;SS go:PFV-REL
 ‘Crying and crying he went.’

Mosetén marks a clear distinction between repetitions that happen immediately after another and fast (verb root reduplication), over a longer time and with intervals between the repetitions (suffix *-dyi*), and general repetition without specific reference to time intervals between the repetitions (infix) (Sakel 2004, 266).

- (4.24) Mosetén (Mosetenan; Sakel 2004, 270, 270, 270)

- (a) **kaph-kaph-yi**.
 clap-REDUP-VM.M.SBJ
 ‘He claps his hands (several times, rather fast).’
- (b) **kaph-kaph-dyij**.
 clap-REDUP-ITE.M.SBJ
 ‘He claps (constantly).’
- (c) **ka-^hb-e-^h**.
 clap-ITE-clap-VM-3F.OBJ
 ‘He claps at her (several times, general, transitive).’

In two languages, Wari’ and Kwaza, it can be observed that reduplication of the verb root can take place more than once. It is unknown whether the number of reduplications mirrors the exact number of repetitions, e.g. whether in (4.25b) the lightning happened exactly twice.

- (4.25) Kwaza (unclassified; Van der Voort 2004, 457)

- (a) **aru=o’ne=aru= o’ne=aru= o’ne-da-ki**
 cross=come=cross=come=cross=come-1SG-DECL
 ‘I’m jumping to and fro.’
- (b) **wero=we’ro-tse**
 lightning=lightning-DECL
 ‘It is lightning.’

- (4.26) Wari’ (Chapacuran; Everett & Kern 1997, 316)

To’ ’ac xucucun na, **to’ to’ to’ to’**, nama.
 hit travel REFL.3PL.M 3SG.REA.PST/PRS hit hit hit hit stop
 ‘Then they hit each other, they hit (each other) repeatedly (or kept on hitting each other), and stopped.’

ITERATIVE is a category which shows a relatively high degree of alternative ways of marking the same function. Eight languages in the sample exhibit reduplication, but also additional markers for ITERATIVE. Little is known about whether in these languages they really are semantic alternatives or exhibit some functional differences, if they have to co-occur or actually exclude each other. In Aguaruna at least, the ITERATIVE suffix *-kawa/ -kua* always co-occurs with reduplication (Overall 2007, 384) (see (4.23) above).

It is to be expected that when two markers for ITERATIVE occur in one language, they do not have exactly the same function, but that their diachronic development is discriminative. Bybee et al. (1994) comment upon possible grammaticalization paths for reduplication, but do not mention the co-existence of reduplication and other ITERATIVE marking in one language. Their observation that ITERATIVE is mostly marked by reduplication and affixation (ibid. 161) can be confirmed for the present study.

Table 4.4: Formal marking of ITERATIVE in the SAILS

(partial) reduplication of verb	Affixes, particles, clitics	Both
Tsafiki, Wari', Hup, Nasa Yuwe, Yaminahua, Cubeo, Karo, Nheengatú, Emérillon, Kamaiurá, Tapiete, Leko, Yurakaré, Kwaza, Trumai	Baure, Tariana, Aymara, Tehuelche, Rikbaktsa, Mamaindê, Sabanê, Warao	Aguaruna, Mosestén, Matses, Huallaga Quechua, Cavineña, Mekens (?), Cocama-Cocamilla, Itonama

4.6 COMPLETIVE/ INCOMPLETIVE

4.6.1 Definitions

COMPLETIVE and INCOMPLETIVE depict very specific Aspectual meanings with many cross-overs to the other Aspects and also Modality and Tense, so it is necessary to have especially fine-grained definitions. The following features characterize a COMPLETIVE marker:

1. The emphasis is on an action that is finished at the point of reference.
2. It is not a PERFECTIVE (i.e. regarding the action as a whole without focusing on the termination).
3. May include a meaning of an action being done completely, wholly, thoroughly.

A COMPLETIVE marker refers to an action whose focus is that is finished or terminated, and may also or only mean that the termination was done completely. Although these meanings are sometimes differentiated by different markers, in the present study they are considered as one. COMPLETIVE is not to be confounded with Aktionsart verbs which have an inherent endpoint (telic). With a COMPLETIVE the speaker stresses that the action has been finished, and even finished thoroughly: “to do something thoroughly and to completion, e.g. *to shoot someone dead, to eat up*” (Bybee et al. 1994, 318). It is also easily confounded with PERFECTIVE, but “the use of the perfective puts no more emphasis, necessarily, on the end of a situation than on any other part of the situation” (Palmer 2001, 18). Like the PERFECTIVE, the COMPLETIVE can have an internal structure, e.g. the COMPLETIVE suffix *-pa* in Cocama-Cocamilla (Vallejos

Yopán 2010, 363): *erutsu-ka-pa* ‘carry something away completely’. The COMPLETIVE can apply to events that just started, so although the event is still in progress, the starting of it is complete, like in Awa Pit:

(4.27) Awa Pit (Barbacoan; Curnow 1997, 231)

ma=na alu ki-ma-ti
 now=TOP rain rain-COMPL-TERM
 ‘It has just started raining.’

The following features characterize an INCOMPLETIVE:

1. Emphasizes that an action is either not finished or not started (yet)
2. NOT CONTINUATIVE
3. NOT FRUSTRATIVE

An INCOMPLETIVE marker is the opposite of COMPLETIVE and denotes emphasis that an action is not (yet) completed or never started. It is not to be confused with CONTINUATIVE or FRUSTRATIVE, though they share certain semantic components. A COMPLETIVE automatically entails that the action is unfinished, but being unfinished does not entail that the action is still in progress (i.e. CONTINUATIVE); although the INCOMPLETIVE can mark actions that are ongoing/ true at the moment of speaking. It is typically translated with ‘still’.

(4.28) Hup (Makuan; Epps 2008, 585)

(a) *pib tɛ, ʔáh-áh*
 strong INCP 1SG-DECL

‘I’m still strong.’

(b) *b’oy-nth tɛ*
 study-NEG INCP

‘(He’s) not studying yet’ ~ ‘hasn’t gone to school yet.’

INCOMPLETIVE does not entail that the action is FRUSTRATIVE; though it may mean that an action was not carried out effectively it does not have the additional element of a desired or expected outcome. It often has a connotation that the action can still be finished, which is not the case with FRUSTRATIVE. A marker can be vague about being FRUSTRATIVE or INCOMPLETIVE, as in Matses and Kwaza. The Matses suffix *-uid* can refer to both an action that was not finished, and an action that was not finished and additionally was expected to have a different outcome. In Kwaza, suffix *-le* can have the meaning of an unfinished but desired action, but also a sense of unrealizedness (‘nearly’) and even progressivity (Van der Voort 2004, 432, 433). In this case, *-uid* and *-le* are coded as vague between INCOMPLETIVE and FRUSTRATIVE.

(4.29) Matses (Panoan; Fleck 2003, 362)

- (a) *cun tied neshca-uid-o-mbi*
 1GEN swidden weed-INCP.FRUST-PST-1SG
 ‘I started weeding my swidden but did not quite finish.’
- (b) *shēctenamē cues-uid-o-mbi*
 white.lipped.peccary kill-INCP.FRST-PAST-1SG
 ‘I ineffectively tried to kill a peccary. [i.e. wounded it, but it escaped]’

4.6.2 Distribution in the sample

Of the 63 languages in the sample, 26 morpho-syntactically mark *COMPLETIVE* and 14 *INCOMPLETIVE*. Of these, eight languages mark both *COMPLETIVE* and *INCOMPLETIVE*, 19 mark *COMPLETIVE*, but not *INCOMPLETIVE*, and six languages mark *INCOMPLETIVE*, but not *COMPLETIVE* (see table 4.5).

One peculiarity that catches the eye is that the *COMPLETIVE* meaning can subtly change from affecting a single referent completely to affecting all of the multiple referents, as in e.g. Cavineña: Cavineña marks *COMPLETIVE* by the suffixes *-tere* (transitive) and *-tirya* (intransitive) (Guillaume 2008, 191–192), referring to an action that was done completely. The meaning is also affected by the number of the referents; when it is singular, it marks that the whole referent is affected, when it is plural, all the referents are affected (ibid. 193).

(4.30) Cavineña (Tacanan, Guillaume 2008, 192, 193)

- (a) *tiru-tere-wa e-kwe-budari*
 burn-COMPL-PFV 1SG-GEN-banana
 ‘My banana burned completely.’
- (b) *iji=ra=e-kwe maju-tere-jeri-kware*
 drink=CAUS=1SG-DAT die-COMPL-ALMOST-REM.PST
 ‘(Because of the drought,) my horses have almost all died from thirst.’

In Dâw, the *COMPLETIVE* marker *hũ?* (which derives from the verb for ‘finish, complete’ (Martins 2004, 290)) together with multiple referents refers to the totality of referents rather than total accomplishment of the action (4.31c).

(4.31) Dâw (Nadahup; Martins 2004, 291, 292)

- (a) *tih ne hũ? cem*
 3SG make COMPL night
 ‘He made everything at night.’ (orig. ‘Ele fez tudo á noite.’)
- (b) *tih xrd hũ? fãmãh tɔp but*
 3SG seek COMPL Xamã house in
 ‘He searched the whole house for Xamã.’ (orig. ‘Ele procurou tudo na casa do Xamã.’)

- (c) *wám fun ʔox hām hūʔ tih wap*
 cutiporó COL run go COMPL 3SG TOT.AUG
 ‘The cutiporós escaped, all of them.’ (orig. ‘Os cutiporós fugiram, todos eles.’)

Cavineña and Dâw mark both COMPLETIVE and INCOMPLETIVE, but the same phenomenon can also be observed in languages marking only COMPLETIVE, mostly in Tupían (Guaraní) (i.e. Emérillon, Kamaiurá, Cocama-Cocamilla), but also in Desano (and probably in the other COMPLETIVE marking languages as well, but the lack of data inhibits to state this with certainty). All of these make a distinction in meaning according to the verbal valency: The Emérillon COMPLETIVE suffix *-pa/ -ba* indicates that the action is completely realized by all the referents of the subject of an intransitive verb, or with respect to all the referents of the object of a transitive verb (Rose 2003, 421).

(4.32) Emérillon (Tupían, Guaraní; Rose 2003, 422)

- (a) *o-ho-pa-ŋ*
 3-go-COMPL-PL.S
 ‘They all have left.’ (orig. ‘Ils sont tous partis.’)
- (b) *o-ʔu-pa*
 3-eat-COMPL
 ‘He ate all (of it).’ (orig. ‘Il a tout mangé.’)

Likewise, the Kamaiurá COMPLETIVE suffix *-pap* refers to the totality of all the objects with transitive verbs, and to the totality of the subjects with intransitive verbs (Seki 2000, 134).³

(4.33) Kamaiurá (Tupían, Guaraní; Seki 2000, 134)

- (a) *o-juka-pap*
 3-kill-COMPL
 ‘He killed all.’ (orig. ‘ele matou todos.’)
- (b) *o-mano-pap*
 3-die-COMPL
 ‘They all died.’ (orig. ‘morreram todos.’)

Cocama-Cocamilla marks COMPLETIVE with suffix *-pa*. In transitive clauses, all objects are affected, in intransitive ones, all subjects are affected.

³ According to Aikhenvald (2012, 180), there is another COMPLETIVE marker in Kamaiurá: suffix *-katu* which grammaticalized from the stative root meaning ‘be.good’.

(4.34) Cocama-Cocamilla (Tupían, Guaraní; Vallejos Yopán 2010, 362, 365)

- (a) *inu-eyu-pa* *ya=pura*
 3PL.FEMALESPEECH-eat-COMPL 3SG.FEMALESPEECH=FOC
 ‘They eat it up.’
- (b) *aypa-pa* *raepe rana*
 grow.up-COMPL then 3PL.MALESPEECH
 ‘Then they finish growing up.’

In Hixkaryana, the COMPLETIVE suffix *-tihka* can mean that emphasis is on the finishing of the action or that all of the referents were affected:

(4.35) Hixkaryana (Cariban; Derbyshire 1985, 225)

- ni-ahataka-tihka-txowni*
 3SBJ-come.out-COMPL-DIST.PST. PFV.COLL
 ‘They finished coming.’ or ‘They all came out.’

Mamaindê and Desano have different COMPLETIVE markers that distinguish between a finished action and an action that was done completely. The Mamaindê marker *-hã?*/ *-hãn* marks an action that was done completely, and *-talona* an action that is finished. They can co-occur:

(4.36) Mamaindê (Nambikwaran; Eberhard 2009, 409)

- (a) *jain-hã?-ten-a?-ø-wa*
 eat-COMPL.all-DES-1SG-PRES-DECL
 ‘I will eat it all.’
- (b) *jalik haʔĩin wek-talona-ø-t^hunna-wa*
 necklace quickly make-COMPL.finish-3SG-FUT-DECL
 ‘She will finish the necklace quickly.’
- (c) *jalik-nãʔã haʔĩin wek-talona-hã?-ø-t^hunna-wa*
 necklace-PL quickly make-COMPL.finish-COMPL.all-3SG-FUT-DECL
 ‘She will finish all the necklaces quickly.’

One would expect that COMPLETIVE most often occurs with PAST reference, as an action which is finished is necessarily situated before the point of speech. That this is not the general rule, however, is demonstrated by the Hup COMPLETIVE suffix *-cĩp/ -cĩw*, which can co-occur with FUTURE marking and both PRESENT and PAST time reference:

(4.37) Hup (Nadahup; Epps 2008, 551, 550, 553)

- (a) *ʔãh j'əm-hi-c̣ip-té-h*
 1SG bathe-FCT-COMPL-FUT-DECL
 'I'll finish bathing.'
- (b) *ʔédia hipãh-c̣iw-ty ʔühny*
 Elias know-COMPL-DYN maybe
 'Elias already knows, maybe.'
- (c) *tedé=d'əh-ət tih bi?ni-c̣ip-ih*
 three=PL-OBQ 3SG work-be-COMPL-DECL
 'He's already worked with three (of them).'

Several COMPLETIVE markers with the meaning of finishing or terminating an action derive from lexemes with similar lexical content, as in e.g. Hixkaryana and Tariana. The Hixkaryana COMPLETIVE suffix *-tihka* also occurs as verbal stem with the meaning 'to finish' (Derbyshire 1985, 225). The Tariana enclitic =*sita* marks actions that are accomplished and derives from a construction with the grammaticalized verb *sita* 'finish, manage' (Aikhenvald 2003, 337–338).

A diachronic source of an INCOMPLETIVE marker in Dâw is the verb for 'to distance, being distant' (Martins 2004, 288). Although Martins describes it as IMPERFECTIVE, it is more characteristic of INCOMPLETIVE, because it does not have HABITUAL OR CONTINUATIVE meanings. (See (4.38a) for *ta?* as verb, and (4.38b) for *ta?* as INCOMPLETIVE.)

(4.38) Dâw (Nadahup; Martins 2004, 288, 289)

- (a) *tàg ta? nã*
 there.it.is it.is.distant disse
 'There it is; it is distant, he says.' (orig. 'Lá está; está distante, disse.')
- (b) *tih kaʔãm ta? j'ãmxu? xad*
 3SG die INCP jaguar because.of
 'He almost died because of the jaguar.' (orig. 'Ele quase morreu por causa da onça.')

COMPLETIVE often involves emphasis. Several, if not all, of the examples in this section can be interpreted in situations where the speaker put focus on a part (the finishing or doing completely) of an action when he wants to emphasize said part. This cannot be proven as most examples are given out of context, but "there is also a certain emphasis inherent in the notion of having brought an action to a thorough conclusion" (Bybee et al. 1994, 57).

The results from studying COMPLETIVE markers in the SAILS sample show that these share the feature of marking the completeness/ thoroughness of an action or the total affectedness of the subject(s)/object(s), and that this relates to emphasis. This is in line with the results for the study carried out by Bybee et al. (ibid.), who think that "if an action is

done completely, it is likely to affect the object totally and may well involve multiple entities”.

Table 4.5: COMPLETIVE/ INCOMPLETIVE marking in the SAILs

COMPLETIVE only	INCOMPLETIVE only	Both
Ika, Kamaiurá, Awa Pit, Pilagá, Yanesha', Urarina, Rikbaktsa, Emérillon, Cocama-Cocamilla, Paresi, Mamaindê, Miraña, Embera, Hixkaryana, Desano, Sateré-Mawé, Shipibo-Konibo, Baure, Leko	Movima, Matses, Cuzco Quechua, Kwaza, Apurinã, Tiriyó	Aymara, Tariana, Dâw, Hup, Huallaga Quechua, Cavineña, Mosetén, Timbira

4.7 ANTERIOR

The last category in the sample is ANTERIOR, otherwise also known as PERFECT. Because it occurs in only four languages in the sample it is discussed here rather than in a separate section. ANTERIOR/ PERFECT and PERFECTIVE are often used for the same category in the literature, but here ANTERIOR denotes a concept dissimilar to PERFECTIVE. It is closely related to Tense, because an action marked with a ANTERIOR refers to a situation in the past that has relevance for the present. Because this meaning is often not clear from one-line examples and has to be inferred from context, it is so often mixed up with PERFECTIVE. Another reason is that, according to Bybee et al. (1994, 51), ANTERIOR markers are an intermediate grammaticalization stage of auxiliaries to PAST and PERFECTIVE.

I agree with Bybee et al. (1994, 54) that an ANTERIOR marker “signals that the situation occurs prior to reference time and is relevant to the situation at reference time” (cf. also Dahl (1985) and Comrie (1976)), as in (4.39). The speaker relates that he was attacked by a caiman which he previously had cured and was in turn saved by a dog from a caiman. The curing of the caiman is marked with ANTERIOR *-wa* because the caiman was only able to attack him later because the speaker cured him previously (Guillaume 2008, 177).

(4.39) Cavineña (Tacanan; Guillaume 2008, 177)

tu-ra ikwene matuja chachane-wa

3SG.ERG first caiman cure-ANT

amena tume=tu chapa=ra tu-ke

BM then=3SG(-FM) dog=ERG 3SG-FM

‘He (the hunter) had first saved (lit. cured) a caiman (about 10 years earlier). Then, (this time,) a dog saved him.’

The four languages in the sample with overt ANTERIOR marking are Baure, Tariana, Imbabura Quechua, and Cavineña. Other languages may have yet undetected ANTERIOR makers in the disguise of PERFECTIVES or PAST markers, which may explain the rarity of ANTERIOR markers in the sample.

Danielsen (2007, 272) suggests that the ANTERIOR particle *ver* grammaticalized from the adverb ‘already’, and that extensive contact of Baure speakers with Spanish may have been the instigator. The adverbial use of *ver* is difficult to distinguish from the particle.

(4.40) Baure (Arawakan; Danielsen 2007, 273)

ver *nik* *p-aha-šo-wo-o-i'*
 ANT 1SG.eat one-plate-one-COP-NOM-EMP
 ‘I already ate one whole plate.’

There is no information about sources for the Tariana and Cavineña ANTERIOR markers. However, the Tariana marker =*ni* occurs in several Aspect-related and non-Aspect-related forms: COMPLETIVE =*niki* (Aikhenvald 2003, 340), ITERATIVE =*nipe* (ibid. 461), nominalizer =*nipe* (ibid. 330), and has overtones of confirmation when in combination with the PAST VISUAL markers (ibid. 332). A possible source is the verb *ni* ‘do, make’. Although Tariana and Baure are related (Arawakan) the grammaticalization path of the Baure ANTERIOR marker points toward individual development rather than shared ancestry. That ANTERIOR marking in these languages is a result of contact is unlikely; none of them are in a contiguous region, as can be seen on map 4.1 below. ANTERIOR marking in South America is thus a language individual phenomenon, governed by internal grammaticalization processes and not due to genealogical or geographical factors.

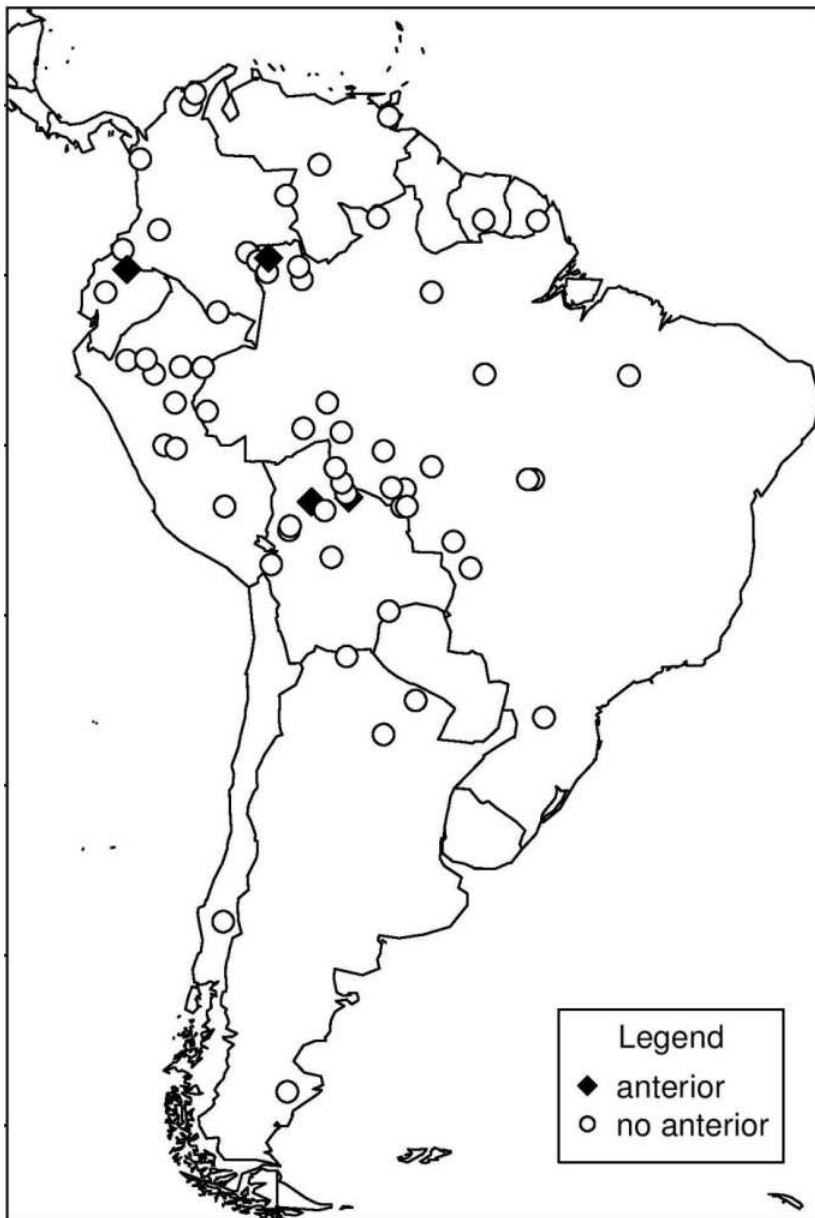
According to Dahl & Velupillai (2011c), there are three languages (of 21) in SA that have ANTERIOR marking: Tucano, Jivaro, and Cocama. ANTERIORS found by Dahl & Velupillai are not derived from possessive constructions or words with the meanings ‘finish’ or ‘already’, whereas at least for Baure the ANTERIOR marker clearly derives from the adverb ‘already’. Because there are some differences between coding in that and the present study the difference is not surprising; however, it is clear that ANTERIOR is a very rare category in SA, regardless of diverging definitions.

Table 4.6: ANTERIOR markers in the sample

Language	Family	ANTERIOR	source
Baure	Arawakan	particle <i>ver</i>	adverb <i>ver</i> ‘already’ (Danielsen 2007, 272)
Tariana	Arawakan	enclitic = <i>nhi</i>	?
Imbabura Quechua	Quechuan	suffix = <i>shka</i>	?
Cavineña	Tacanan	suffix = <i>wa</i>	?

Map 4.1: Distribution of ANTERIOR marking in the sample

Anterior



4.8 TYPOLOGICAL DISTRIBUTION IN THE SAMPLE

Of the 63 languages in the sample, no language is completely without Aspect marking. Six languages mark one Aspectual, 15 mark two, 23 mark three, 12 mark four, and seven mark five. The most frequently marked category is CONTINUATIVE (35), followed by ITERATIVE (32) and HABITUAL (31), COMPLETIVE (26), PERFECTIVE (24), IMPERFECTIVE (22), INCOMPLETIVE (14), and ANTERIOR (4). Every language in the sample has either IMPERFECTIVE or one of its subcategories marked, but 19 languages do not mark PERFECT OR COMPLETIVE. This section investigates several points of interest that arise from the results, starting with a comparison with previous studies, before coming to the genealogical and geographical analysis. As a last point, the stability of certain Aspectuals is discussed.

In the global sample of Dahl (1985, 183-185) the CONTINUATIVE (“progressive”) occurs less frequently than PERFECTIVE/ IMPERFECTIVE and is predominantly expressed by periphrastic means. Only in one out of 19 instances is CONTINUATIVE marked by inflection. In contrast, CONTINUATIVE is the most frequently marked category in the present sample and predominantly marked by inflection (23 instances of suffixation, four of cliticization, and one of prefixation). There is no comparison possible as to whether the number of periphrastic CONTINUATIVE marking is higher than that of inflectional marking, but the frequency of CONTINUATIVE inflectional marking and CONTINUATIVE in comparison with PERFECTIVE/ IMPERFECTIVE is significantly higher than in Dahl’s sample. The fact that most languages in this study morpho-syntactically mark CONTINUATIVE is in line with Bybee et al.’s (1994, 174) results, where “[t]he most widespread and frequently occurring [...] is progressive, which is usually periphrastic in expression”. It is not possible to say whether the CONTINUATIVE in the SAILS is more frequently expressed periphrastically than morpho-syntactically. But we can assume that the SAILS do express CONTINUATIVE with an unusually high number of morpho-syntactic markers compared to other regions of the world. It was already noted in 4.2.2 that PERFECTIVE/ IMPERFECTIVE marking occurs in about half of the present sample as well as in a global sample.

The rather high frequency of HABITUAL markers in the sample also deviates from Dahl’s (1985, 95-96) results, which not only show a low frequency of forms with HABITUAL, HABITUAL GENERIC, and HABITUAL PAST meaning, but are also mostly expressed periphrastically. A low frequency of HABITUAL also occurs in Bybee et al. (1994, 160), although their study even includes periphrastic constructions. This points to the following generalization: SAILS have a higher frequency of marking HABITUAL compared to Dahl’s and Bybee et al.’s samples. This in turn leads to the conclusion that frequent HABITUAL marking is a characteristic feature of the SAILS in the sample.

Figure 4.2: Aspectuals marked by SAILs

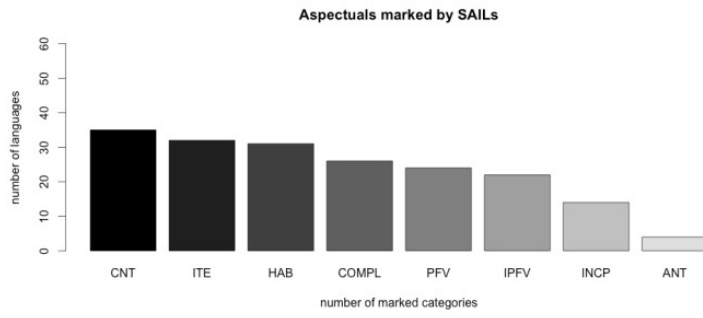


Figure 4.3: Number of SAILs marking number of Aspectuals

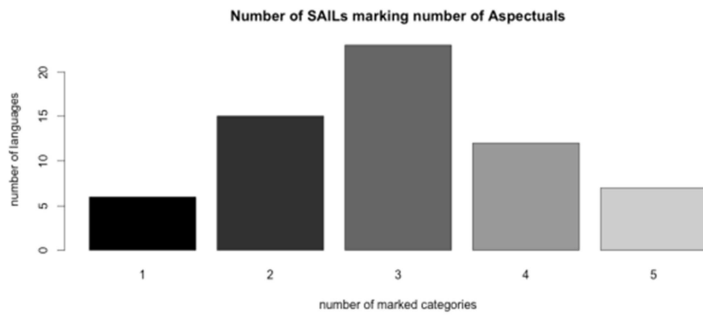


Table 4.12: Aspect marking in the sample

	LANGUAGE	FAMILY	CNT	ITE	HAB	COMPL	PFV	IPFV	INCP	ANT	TOTAL
1	Mocoví	Guaycuruan	■								1
2	Karitiana	Tupían	■								1
3	Karo	Tupían		■							1
4	Kanoê	Unclassified						■			1
5	Wichí (Mataco)	Matacoan						■			1
6	Movima	Unclassified							■		1
7	Tsafiki	Barbacoan	■	■							2
8	Wari'	Chapacuran	■	■							2
9	Itonama	Unclassified	■	■							2
10	Tehuelche	Chonan	■	■							2
11	Sabanê	Nambikwaran	■	■							2
12	Jarawara	Arawan	■		■						2
13	Mapuche	Araucanian	■		■						2
14	Ika	Chibchan, Aruak	■			■					2
15	Tapiete	Tupían, Guaraní		■	■						2
16	Kamaiurá	Tupían, Guaraní		■		■					2
17	Aguaruna	Jivaroan		■				■			2
18	Matses	Panoan		■					■		2
19	Awa Pit	Barbacoan				■		■			2
20	Chimila	Chibchan					■	■			2
21	Yanam	Yanomaman					■	■			2
22	Bororo	Macro-Gêan	■	■	■						3
23	Mekens	Tupían	■	■	■						3
24	Cubeo	Tucanoan	■	■	■						3
25	Rikbaktsa	Macro-Gêan	■	■		■					3
26	Emérillon	Tupían, Guaraní	■	■		■					3
27	Cocama-Cocamilla	Tupían, Guaraní	■	■		■					3
28	Mosetén	Mosetenan	■	■		■			■		3
29	Pilagá	Guaycuruan	■		■	■					3
30	Yanesha'	Arawakan	■		■	■					3
31	Urarina	Unclassified	■		■	■					3
32	Puinave	Unclassified	■		■			■			3

33	Paresi	Arawakan									3
34	Timbira	Macro-Géan									3
35	Munichi	Unclassified									3
36	Cuzco Quechua	Quechuan									3
37	Mamaïndê	Nambikwaran									3
38	Warao	Unclassified									3
39	Trumai	Unclassified									3
40	Kwaza	Unclassified									3
41	Miraña	Boran									3
42	Embera	Chocoan									3
43	Kaingang	Macro-Géan									3
44	Hixkaryana	Cariban									3
45	Nasa Yuwe	Paezan									4
46	Aymara	Aymaran									4
47	Desano	Tucanoan									4
48	Sateré-Mawé	Tupían									4
49	Panare	Cariban									4
50	Apurinã	Arawakan									4
51	Imbabura Quechua	Quechuan									4
52	Nheengatú	Tupían, Guaraní									4
53	Yurakaré	Unclassified									4
54	Yaminahua	Panoan									4
55	Shipibo-Konibo	Panoan									4
56	Tiriyó	Cariban									4
57	Leko	Unclassified									5
58	Baure	Arawakan									5
59	Dâw	Nadahup									5
60	Tariana	Arawakan									5
61	Hup	Nadahup									5
62	Huallaga Quechua	Quechuan									5
63	Cavineña	Tacanan									5
			35	32	31	26	24	22	14	4	

4.9 GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE

This section focuses on geographical patterns of Aspect marking in the sample. There are a few noteworthy geographical patterns of Aspect marking, although mostly no clear picture emerges. In the languages of the Guaporé-Mamoré and adjacent regions mostly neither PERFECTIVE nor IMPERFECTIVE are marked. Languages which mark both are found clustering in Peru/ western Brazil (Munichi, Yaminahua, Shipibo-Konibo, Huallaga Quechua), northern Brazil to Surinam (Yanam, Hixkaryana, Tiriyo), and in single instances in southern Brazil (Kaingang), north-west Brazil (Nheengatú), the Andean fringe (Yurakaré, Leko), Colombia (Chimila) and Venezuela (Panare). Of these, the cluster in Peru/ Brazil is most interesting, because all these languages are in a contiguous region and have reportedly been in contact. The PERFECTIVE/ IMPERFECTIVE forms of Huallaga Quechua are not cognate with those of the other Quechuan languages in the sample and neither do those have a full PERFECTIVE/ IMPERFECTIVE distinction. Shipibo-Konibo and Yaminahua are both Panoan languages which suggests that their Aspectuals are rather cognates than contact phenomena, although a third Panoan language, Matses, does not mark either PERFECTIVE or IMPERFECTIVE. Munichi is still unclassified.

There is a similar distribution of COMPLETIVE/ INCOMPLETIVE marking. Marking of both categories in one language occurs only in isolated instances throughout the continent, but an accumulation of marking of COMPLETIVE occurs in approximately the same region as the PERFECTIVE/ IMPERFECTIVE marking cluster above: Shipibo-Konibo, Huallaga Quechua, Yanasha', Urarina, Cocama-Cocamilla.

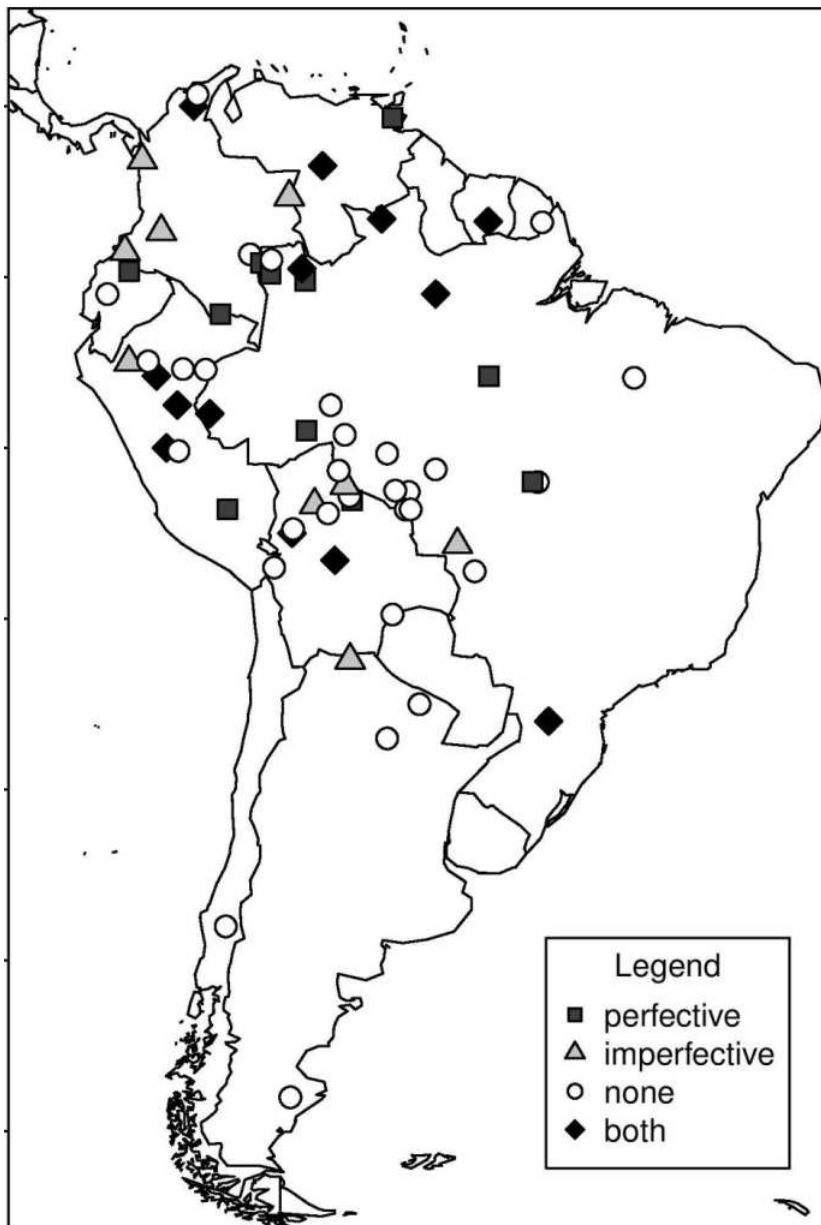
ITERATIVE marking occurs mostly in three regional clusters: around and including the Vaupés area, Bolivia/ Rondônia, and northern Peru, although all of these regions also feature languages without ITERATIVE marking. Interestingly, ITERATIVE marking is noticeably absent in Amazonia, although Rubino (2011) claims that reduplication in general is a very common phenomenon in Amazonia. However, his data is not convincing: first, he does not specify "Amazonia" and second, the corresponding WALS map does not show any particular presence of reduplication in any region that could be termed 'Amazonia'. Additionally, the number of languages presented in that region is rather low (although the same case could be made for this study). It is possible, though, that reduplication as ITERATIVE marking is absent in a language but fully present in other parts of grammar, so Rubino's and the present data do not necessarily have to contradict each other. What Rubino does find, but does not comment upon, is an accumulation of reduplication, both full and partial, in Bolivia and Rondônia, which agrees with the occurrence of reduplication in this study. Map 4.3 below illustrates that most of the languages in this sample in that region mark ITERATIVE either only by reduplication or by both reduplication and other means (affixes, clitics, or particles). This is not surprising, because when reduplication is a productive grammatical device in a language in general, it is likely that it is available for ITERATIVE marking as well.

Another noticeable absence of marking is that of HABITUAL in the area and adjacent regions of the Guaporé-Mamoré. As was seen above, IMPERFECTIVE (and PERFECTIVE) are also rarely marked in that region, but CONTINUATIVE and ITERATIVE fare slightly better.

To summarize, Aspectual categories or their absence apparently cluster in a few cases in the Vaupés and Guaporé-Mamoré/ Bolivia and Rondônia, and northern Peru.

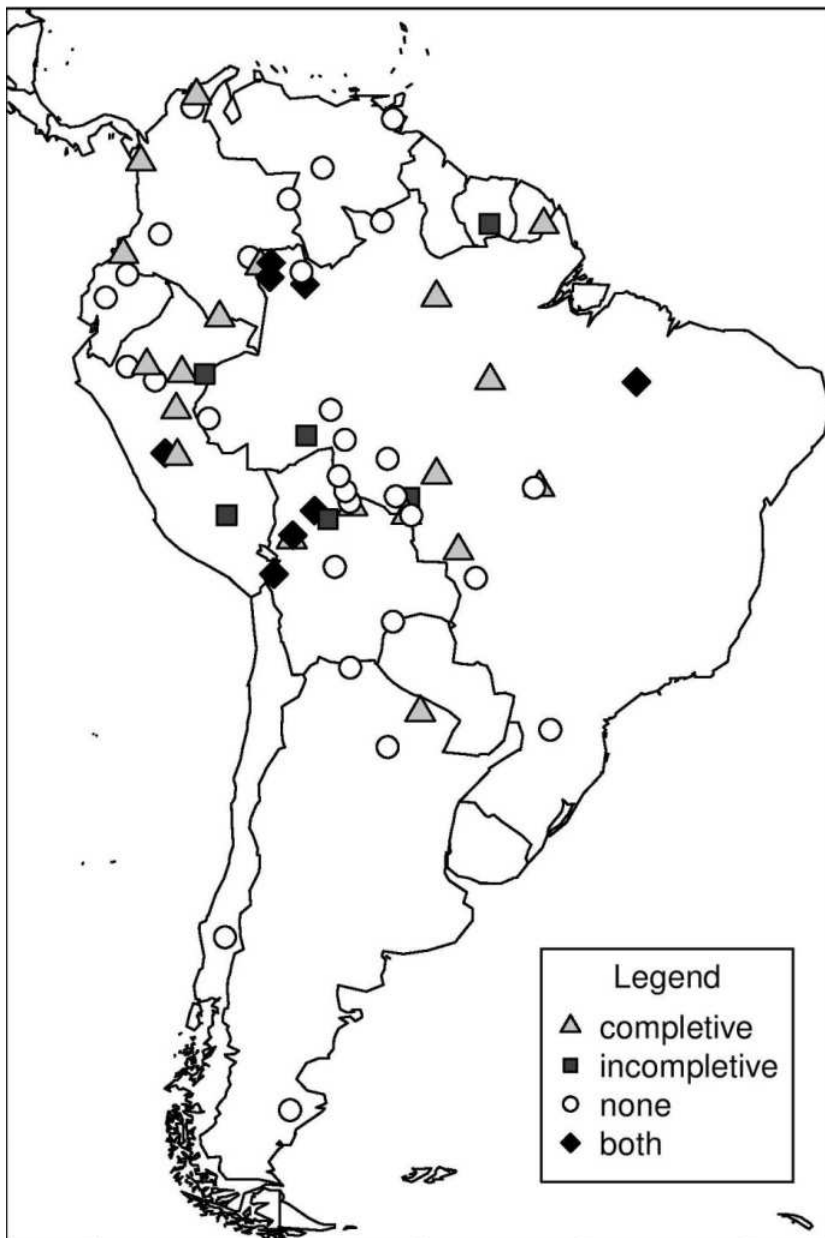
Map 4.2: Distribution of PERFECTIVE/ IMPERFECTIVE marking in the sample

Perfective/ Imperfective marking



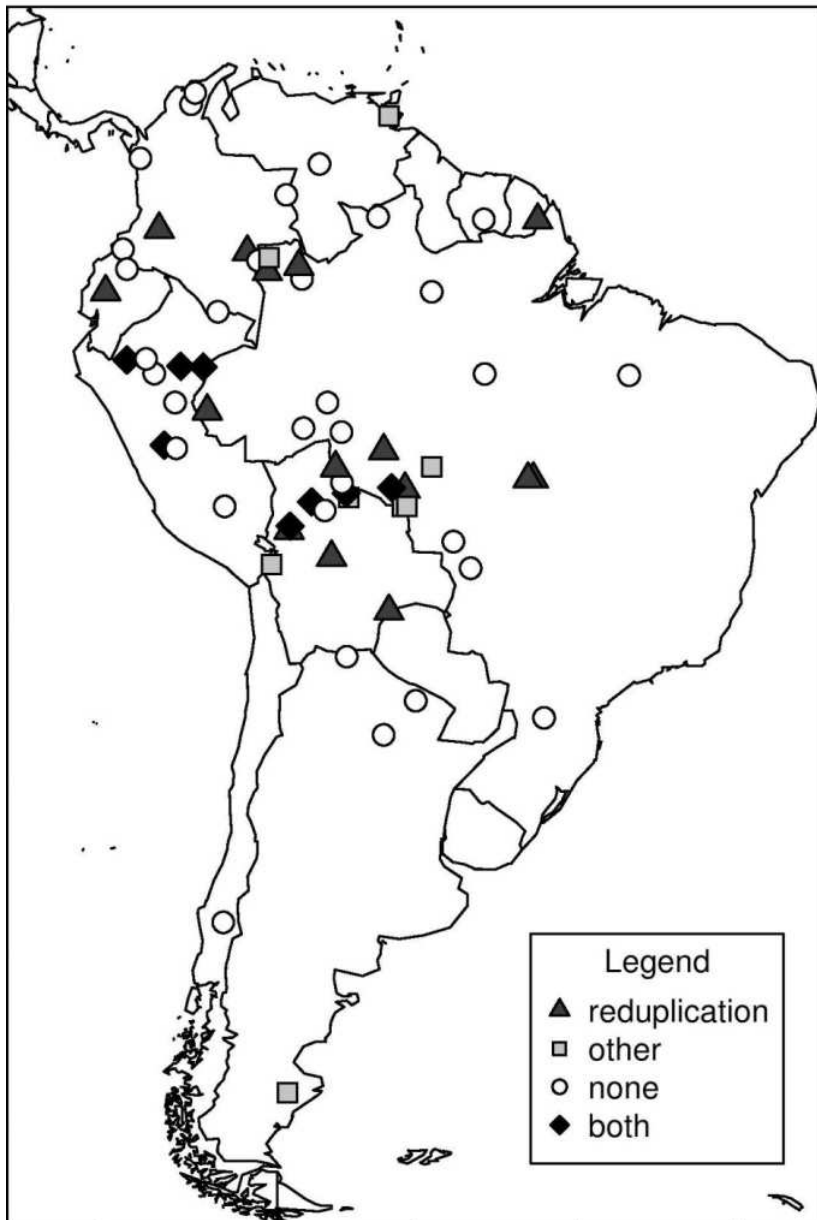
Map 4.3: Distribution of COMPLETEIVE/ INCOMPLETEIVE marking in the sample

Completeive/ Incompleteive marking



Map 4.4: Distribution of ITERATIVE marking in SA

Iterative marking types



4.10 GENEALOGICAL DISTRIBUTION IN THE SAMPLE

4.10.1 Introduction

The following paragraphs present investigations of Aspectual systems in selected language families: Macro-Gêan (4.10.2), Arawakan (4.10.3), Tupían (4.10.4), Cariban (4.10.5), and Quechuan (4.10.6).

4.10.2 Macro-Gêan

Aspect marking in the Macro-Gêan languages in the sample is highly heterogeneous. There is no formal similarity of the markers and very little consistency in marking of the categories. CONTINUATIVE is marked by three languages, HABITUAL, COMPLETIVE, and ITERATIVE by two, and PERFECTIVE, IMPERFECTIVE, and INCOMPLETEIVE only by one. None of the forms are cognates. Rodrigues (1999, 188-189) lists a suffix *ma-* for IMPERFECTIVE in Guató, and a preposition *kri* for PERFECTIVE which also do not seem to be cognates.

Table 4.13: Aspect marking in Macro-Gêan

	Bororo	Rikbaktsa	Kaingang	Timbira
PFV	-	-	<i>mũ</i>	-
IPV	-	-	<i>tí</i>	-
HAB	<i>=kigodũ</i>	-	<i>fã, kamã</i>	-
CNT	<i>=nü</i>	<i>-kv(C)</i>	-	<i>apu</i>
ITE	Redup	<i>-ɽo</i>	-	-
COMPL	-	<i>-ba</i>	-	<i>ramã</i>
INCP	-	-	-	<i>kormã</i>

4.10.3 Arawakan

The Arawakan languages in the sample vary according to which Aspectual category they mark, but exhibit a certain amount of similarity in formal marking. Payne (1991, 381), in his attempt to reconstruct Proto-Maipuran (Maipuran being another name for Arawakan), reconstructs **-pe* for PAST, PERFECTIVE, and COMPLETIVE. This proto-form apparently survived in Apurinã and Baure PERFECTIVE marking, but not in Tariana, Yanasha', and Paresi. Paresi also features the IMPERFECTIVE particle *hena*, which, according to Payne (1991, 381), stems from the Proto-Maipuran form **-ena*. Paresi is the only Arawakan language marking IMPERFECTIVE. Baure and Paresi have similar CONTINUATIVE markers: particle *ito* and suffix *-ita*, respectively, which do not concur with the proto-forms given by Payne. This may be a form common to the South-Arawak branch only.

Besides the formal similarities in PERFECTIVE (Baure and Apurinã) and CONTINUATIVE (Baure and Paresi) marking there is no consistency in Arawakan Aspect marking in the sample. Derbyshire (1986, 518-520) mentions COMPLETIVE suffixes for several Brazilian Arawakan languages (see table 4.10), but they are very heterogeneous with the exception of Waura and Palikur which feature a suffix *-wi* for PERFECTIVE OR COMPLETIVE. The element *-nã, -na, -ne, -n*, which occurs in the CONTINUATIVE forms in Apurinã, Yanasha', Dení, Palikur, Amuesha, and in the IMPERFECTIVE in Paresi is possible a reflex of the proto-form suggested by Payne

(1991), but this needs further research. The heterogeneity of the Aspectuals is also mirrored in the PreAndine Arawakan language groups Campa and Amuesha given by Wise (1986, 587-588): Campa has a PERFECTIVE OR COMPLETIVE suffix *-ak* and a PROGRESSIVE suffix *-ač*, and Amuesha has a PROGRESSIVE suffix *-en* and a COMPLETIVE suffix *-om*. The Campa PERFECTIVE suffix may, however, rather be an ANTERIOR (due to a very common confusion between the terms PERFECTIVE and PERFECT).

Aspectuals in Arawakan are only partially family specific: the PERFECTIVE forms in Baure and Apurinā, the IMPERFECTIVE in Paresi, the COMPLETIVE OR PERFECTIVE in Waura and Palikur, and possibly the CONTINUATIVE OR IMPERFECTIVE in Apurinā, Yaneshá', Dení, Palikur, Amuesha, and Paresi.

Table 4.14: Aspect marking in Arawakan

	PFV	IPFV	HAB	CNT	ITE	COMPL	INCP
Proto-forms ⁴	* <i>pe</i>	* <i>ki</i> , * <i>ena</i>	* <i>ki</i> , * <i>ena</i>	* <i>ki</i> , * <i>ena</i>	-	* <i>pe</i>	-
Baure	<i>-po</i>	-	-	<i>ito</i>	<i>avik</i> , <i>-poreiy</i>	<i>eto</i>	-
Paresi	-	<i>hena</i>	-	<i>-ita</i>	-	<i>heta</i>	-
Apurinā	<i>-pe</i>	-	<i>-pi</i>	<i>nanu</i>	-	-	<i>-panhi</i>
Yaneshá'	-	-	= <i>e't̃</i>	= <i>meñ</i>	-	<i>-Vhu/ -V'hu/ -V'hua/ -u/ -ua</i>	-
Tariana	-	-	= <i>hyuna</i> , = <i>kape</i> , = <i>nipe</i>	-	<i>-nipe</i>	= <i>niki</i> , = <i>sita</i>	= <i>daka</i> , = <i>sida</i>
Dení ⁵	?	?	?	<i>-nava</i>	?	<i>-ni/ -vi</i>	?
Palikur	?	?	?	<i>-n/ -ne(ne)</i> , <i>-(na)no</i>	?	<i>-e/ -i/ -wi/ -pi/ -ep</i>	?
Terêna	?	?	?	<i>-ti</i>	?	?	?
Paumarí	?	?	?	?	?	<i>-i/ -a</i>	?
Waura	<i>-wi</i>	?	<i>-pai</i>	?	?	<i>-ene</i>	?
Campa ⁶	<i>-ak</i>	?	?	<i>-ač</i>	?	<i>-ak</i>	?
Amuesha	?	?	?	<i>-en</i>	?	<i>-om</i>	?

4.10.4 Tupían

With one exception, COMPLETIVE marking in Guaraní, there is no formal similarity among the Aspectuals in the Tupían languages of the sample. Most of the Tupían languages mark about two Aspectuals. Sateré-Mawé and Nheengatú mark the most categories (four), Karitiana and Karo one. The most frequently marked category is ITERATIVE (seven), followed by CONTINUATIVE (five); the least marked is IMPERFECTIVE.

All Guaraní languages except for Tapiete mark COMPLETIVE with cognates which are reflexes of the proto-COMPLETIVE suggested by Jensen (1998, 537): **-pab*,⁷ that with

⁴ Payne (1991, 381)

⁵ Derbyshire (1986, 518-520)

⁶ Wise (1986, 587-588)

intransitive verbs indicates that all subjects perform the action and with transitive verbs indicates that the action is performed on all objects. There is no COMPLETIVE marker in Tapiete, but the RESULTATIVE suffix *-ma* could be a cognate (cf. González 2005, 157). Interestingly, Rikbaktsa, which has been classified as Macro-Gêan (Rodrigues 1999, 168), also has a COMPLETIVE marker of a similar form (*-ba*).

The Tapiete HABITUAL suffix is an illustrative example of language internal development: it grammaticalized from the adverb *yepi* ‘always’ (González 2005, 158).

Table 4.15: Aspect marking in Tupían

	PFV	IPFV	HAB	CNT	ITE	COMPL	INCP
Karitiana	-	-	-	<i>-tiso, -tika,</i> <i>-tijnã,</i> <i>-tisip, -gi</i>	-	-	-
Karo	-	-	-	-	Redup	-	-
Mekens	-	-	<i>kakwat</i>	Auxiliaries	Redup, <i>neara</i>	-	-
Sateré-Mawé	<i>ra'in/</i> <i>ta'in/</i> <i>na'in</i>	-	<i>in</i>	Redup, <i>te</i>	-	<i>ine</i>	-
Nheengatú (Guaraní)	<i>=wã</i>	<i>=re</i>	<i>wera</i>	-	Redup	-	-
Emérillon (Guaraní)	-	-	-	<i>=o, -(i)ŋ</i>	Redup	<i>-pa/ -ba</i>	-
Kamaiurá (Guaraní)	-	-	-	-	Redup	<i>-pap, -katu</i>	-
Cocama- Cocamilla (Guaraní)	-	-	-	<i>-(a)ri</i>	Redup <i>-ka</i>	<i>-pa</i>	-
Tapiete (Guaraní)	-	-	<i>-pi</i>	-	Redup	-	-

4.10.5 Cariban

All three Cariban languages in the sample mark PERFECTIVE and IMPERFECTIVE, but whereas Hixkaryana and Tiriyó both have distinct Aspectual marking of PERFECTIVE/ IMPERFECTIVE according to Tense, Panare apparently has Tense-independent markers. Tiriyó has PERFECT/ IMPERFECTIVE distinctive marking in FUTURE, PRESENT, and PAST, with overt marking all but PRESENT PERFECTIVE (which may be a zero-marker (cf. Meira 1999, 303). Hixkaryana overtly marks PERFECTIVE/ IMPERFECTIVE only in the PAST Tenses (recent and distant), but not in PRESENT or FUTURE (which are cumulative NONPAST morphemes, see 3.7.2, Cariban).

The suffix *-ne* and varieties are common in Cariban for Tense and Aspect marking (Meira 1999, 306-307). It occurs in the majority of Gildea's (1998, 98) sample under 'distant past' although he does not further comment on Aspectual values. Interestingly, it occurs as

⁷ Jensen (1998, 528) gives a verb *paw* 'finish' in Tupían the proto-form of which could be the origin of the COMPLETIVE.

IMPERFECTIVE marker in Panare and PERFECTIVE is marked by *-yaj*, which may be cognate with Gildea's (ibid.) **-ya* in NONPAST forms. This proto-form also occurs in Hixkaryana PAST PERFECTIVE and PAST IMPERFECTIVE as *-ya* and possibly in Tiriyó FUTURE PERFECTIVE, PRESENT IMPERFECTIVE, and PAST IMPERFECTIVE as *-ja*. If these forms are really reflexes of the NONPAST proto-form it would be interesting to identify its development into PAST and the role its Aspectual meaning plays in the process.

The other Aspectual categories are less pronounced. ITERATIVE is unmarked, COMPLETIVE marked by Hixkaryana, INCOMPLETIVE by Tiriyó, HABITUAL by Tiriyó and Panare, and CONTINUATIVE by Panare only. The rest of the Aspect categories are partially expressed by the PERFECTIVE/ IMPERFECTIVE forms; e.g. IMPERFECTIVE forms can be used to express both CONTINUATIVE and HABITUAL in Tiriyó, although the latter even has its own marker (the PAST IMPERFECTIVE is almost completely replaced by the HABITUAL form today (Meira 1999, 332)).

Table 4.16: Aspect marking in Cariban

	Hixkaryana		Tiriyó		Panare
PFV	PRS.PFV	-	PRS.PFV	-	<i>-yaj</i>
	PST.PFV	<i>-yako,</i> <i>-yatxoko,</i> <i>-ye,</i> <i>-txowni</i>	PST.PFV	<i>-ne</i>	
	FUT.PFV	-	FUT.PFV	<i>-(ja)kě(mi)</i>	
IPFV	PRS.IPFV	-	PRS.IPFV	<i>-(ja)-e,</i> <i>-(ja)-(nĕ)</i>	<i>-ñe</i>
	PST.IPFV	<i>-yaknano,</i> <i>-yatxkenano,</i> <i>-yakoni,</i> <i>-yatxkoni</i>	PST.IPFV	<i>-(ja)kě(ne)</i>	
	FUT.IPFV	-	FUT.IPFV	<i>-ta-e,</i> <i>-ta-(nĕ)</i>	
HAB	-	-	<i>-e</i>	-	<i>-sen</i>
CNT	-	-	-	-	<i>-mpĕj/ -nĕpĕj</i>
ITE	-	-	-	-	-
COMPL	<i>-tihka</i>	-	-	-	-
INCP	-	-	<i>nkĕrĕ</i>	-	-

4.10.6 Quechuan

Aspect marking in Quechuan is relatively heterogeneous. All three languages in the sample mark PERFECTIVE. Huallaga Quechua additionally marks IMPERFECTIVE, ITERATIVE, COMPLETIVE, and INCOMPLETIVE, Imbabura Quechua HABITUAL and CONTINUATIVE, and Cuzco Quechua CONTINUATIVE and INCOMPLETIVE. There is a possibility that Imbabura Quechua also marks INCOMPLETIVE.

Cognates of the PERFECTIVE markers found in Tarma Quechua (*-ra/-ru*) (Adelaar with Muysken 2004, 221, 226) and Pacaraos Quechua (*-rqa/-rqu*) (ibid. 243, 244) could be the PAST

markers discussed in 3.7.2. Adelaar with Muysken (2004, 231) suggest that that the reflexes of **-rqu* encode PERFECTIVE in opposition to a PROGRESSIVE *-ya*. Interestingly, both PAST and PERFECTIVE markers can co-occur in at least Pacaraos Quechua. Either the modern PAST and PERFECTIVE forms grammaticalized from the same proto-form (possibly **-rqu*) or there were two similar proto-forms, one for PAST and one for PERFECTIVE. Given the high degree of similarity of the forms and the fact that PERFECTIVE and PAST are established grammaticalization paths from ANTERIOR forms (Bybee et al. 1994, 81), I assume a single proto-form.

(4.41) Pacaraos Quechua (Quechuan; Adelaar with Muysken 2004, 246)

<i>“kuy-çaw</i>	<i>ka-yka-n”</i>	<i>n’a</i>	<i>say-naw-pa-š</i>
that.over.there-LOC	be-PROG-3SBJ	already	that-CP-GEN-HEA
<i>wamra</i>	<i>rima-rqu-rqa</i>		
child	speak-PFV-PST.3SBJ		

‘“It is over there”, a child had said.’

But ANTERIOR, i.e. a marker encoding present relevance of a PAST action, is not the only candidate for the meaning of a proto-form of modern PAST and PERFECTIVE. Adelaar with Muysken (2004, 231) reconstruct **-rku* and **-rqu* as directional affixes encoding ‘upward’ and ‘inward’ movement, respectively, and it has been shown that the grammaticalization path from directionals into at least Tense markers is not uncommon cross-linguistically.

Whereas the Imbabura PERFECTIVE/ PAST marker is a reflex of **-rqu* the forms in Huallaga and Cuzco Quechua are not. PERFECTIVE in Huallaga Quechua is marked by *-ykU* (Weber 1989, 145) and in Cuzco Quechua by *-ña* although the latter is of a dubious nature (Faller (2002) glosses this marker with ‘discontinuative’ but the few examples point towards PERFECTIVE meaning). Only Huallaga Quechua has a fully developed PERFECTIVE/ IMPERFECTIVE opposition.

Table 4.17: Aspect marking in Quechuan

	Huallaga	Imbabura	Cuzco	Proto-forms (Adelaar with Muysken 2004, 231)
PFV	-ykU	-rka	-ña	*-rqu
IPV	-yka	-	-	*-ĕka
HAB	-	-j	-	?
CNT	-	-ju ⁸	-sha	*-ĕka
ITE	-ykacha/-kacha/-cha, reduplication	-	-	?
COMPL	-kaakU	-	-	?
INCP	-raq	- (-raj?)	-raq	?

4.11 Stability of Aspect

This section evaluates the stability of Aspectual features and draws upon the studies by Wichman & Holman (2009) and Wichmann & Kamholz (2008) (see section 3.10 for an introduction to stability of typological features and the stability of Tense, section 4.11 for Aspect, 5.13 for Modality, 6.6 for Evidentiality, and 7.5 for a final discussion). According to Wichmann & Holman (2009, 45) PERFECTIVE/ IMPERFECTIVE are relatively stable and PERFECT (=ANTERIOR) is unstable. If split into the values given in WALS, ANTERIORS derived of possessive constructions are very stable, ANTERIORS derived from words meaning ‘finish’ or ‘already’ are stable, ANTERIORS derived from other sources are very unstable, and absence of ANTERIOR is unstable. Both the absence and presence of the grammatical marking of a PERFECTIVE/ IMPERFECTIVE distinction are stable (ibid. 57-58). PERFECTIVE and IMPERFECTIVE are therefore expected to remain stable over a large period of time, but ANTERIOR not. However, further investigation of ANTERIOR has to wait until the sources from which the four markers in this study developed are identified. As for the stability of PERFECTIVE/ IMPERFECTIVE in the present sample, there is no clear outcome. In some of the language families the absence of both or either one of them is quite stable (e.g. in Tupían and Macro-Gêan the absence of both and in Cariban the presence of both), in others the values differ (e.g. one Quechuan language marks both and the other two only PERFECTIVE). This suggests that in general PERFECTIVE/ IMPERFECTIVE marking in the sample is of overall medium stability and that is approximately in line with Wichmann & Holman (2009). Interestingly, the most frequently marked Aspectual, CONTINUATIVE, is spread over all language families, and isolates. Given its frequency one would expect that it turns out to be a genealogically stable feature, but the reverse is true. Possibly the low stability contributed to the high frequency in the SAILS.

⁸ Cole (1982, 150) claims that this is cognate with the reflexive morpheme *-ku* in Ecuadorian Quechua. Interestingly, the PERFECTIVE suffix *-po* in Baure can also have reflexive meaning (Danielsen 2007, 262). Adelaar with Muysken (2004, 281) state that the suffix *-si*, which marks progressive, is homophonous with the reflexive marker. That reflexive and aspectual marking seem to be related in at least three unrelated languages deserves attention in future research.

In terms of stability of formal marking of a given feature, there is little homogeneity between the markers of languages of the same family. As could be seen above, there are only a few cases where markers of the same feature are sufficiently similar to assume common ancestry (e.g. *-pe for PERFECTIVE in Arawakan, or COMPLETIVE *-pab in Tupían). Because the results for Tense stability markers are similarly heterogeneous and because the semantics of Tense and Aspect interact, the instability of Tense may entail instability of Aspect. As Wichmann & Holman (2009, 34) argue: “One general factor that may be inferred to influence the stability of a feature is the extent to which the feature is related to other features. [...] if changes in a feature are stimulated by changes in related features, then the relationships would promote change and thus decrease stability”. If one can account for changes in Tense that are stimulated by changes in Aspect and *vice versa*, this would corroborate the proposition. Partially, this has been done by e.g. Bybee et al. (1994) who demonstrated that grammaticalization paths of markers often involve a switch from Tense to Aspect or Aspect to Tense. For example, ANTERIOR markers grammaticalize into PAST markers (ibid. 105), and PROGRESSIVES (CONTINUOUS) into PRESENT markers (ibid. 140). Thus, as soon as a marker grammaticalizes from ANTERIOR to PAST, both the Tense and the Aspect systems of a language are affected and decrease in stability. A change in one feature, on the other hand, may also predict a whole series of changes that follow. For example, when an ANTERIOR changes into PAST, it may be that another change is triggered to provide a new ANTERIOR. If this is true, then the two changes would balance each other⁹ and increase the stability of the presence of a feature (but not of the form). However, Tense and Aspect categories were earlier defined as complex semantic fields, with central and peripheral meanings. When a Tense marker develops into an Aspectual one, it is often because of a switch of the central and peripheral meanings, but that does not necessarily involve a loss of meaning which would be the trigger for a new change. The instability of Tense and Aspect marking (of form) can thus be explained by the relationship between the two.

4.12 SUMMARY

This chapter investigated the typological, geographical, and genealogical patterns of Aspect marking in the sample. The languages were analyzed according to PERFECTIVE, IMPERFECTIVE, ANTERIOR, HABITUAL, ITERATIVE, CONTINUATIVE, COMPLETIVE, and INCOMPLETIVE. It was demonstrated that this study focuses on viewpoint Aspect and not Aktionsart, but acknowledges the intricate interaction between the two. Definitions of the individual categories include varying ranges of characteristics that have to be met in order to distinguish them from other Aspectuals. The most detailed list is that of ITERATIVE which shares many characteristics with e.g. HABITUAL, CONTINUATIVE. The biggest challenge of Aspectuals is in their high degree of interrelations, and how to establish where one category ends and

⁹ cf. Wichmann & Holmann (2009, 34) “if the change in a feature requires changes in other related features, then the relationships would impede change and increase stability”.

another begins. It is felt that the study above represents the best possible attempt based on the available data.

It was found that no language in the sample is without Aspect marking, but also that no language marks the possible maximum of eight Aspectual categories. The most frequently marked category is CONTINUATIVE (35), followed by ITERATIVE (32) and HABITUAL (31), COMPLETIVE (26), PERFECTIVE (24), IMPERFECTIVE (22), INCOMPLETIVE (14), and ANTERIOR (4). CONTINUATIVE and HABITUAL are morpho-syntactically marked more frequently in the SAILS than in other regions of the world, whereas PERFECTIVE/ IMPERFECTIVE are marked by about the same portion in SA.

The marking of Aspect within language families varies from being mildly to completely heterogeneous, although suffixation (and reduplication in the case of ITERATIVE) prevails. A few genealogical traits could be identified for Arawakan (*-*pe* PERFECTIVE/ COMPLETIVE), Tupían (*-*pab* COMPLETIVE), Quechuan (PAST/ COMPLETIVE/ PERFECTIVE *-*rqu*) and Cariban (tendency to mark PERFECTIVE/ IMPERFECTIVE distinctions in PAST, PRESENT, and FUTURE by cumulative markers).

The fact that morpho-syntactic marking of both Aspect and Tense is quite heterogeneous was explained by the relationship between the two. It was found that the presence or absence of PERFECTIVE/ IMPERFECTIVE marking is stable in certain language families, but unstable in others, whereas CONTINUATIVE is highly unstable and distributed over all language families and regions.

5. MODALITY

5.1 INTRODUCTION

This chapter investigates the distribution of morpho-syntactic Modality marking with regard to typological, genealogical, and geographic properties in the sample of SAILS. The study concerns REALIS, IRREALIS, INTENTIONAL, CERTAINTY, DUBITATIVE, DESIDERATIVE, FRUSTRATIVE, ITERATIVE, POTENTIAL, PURPOSIVE, and COMMAND types.

The term ‘Modality’ here refers to the concepts of both mood and modality as found in the literature: “[m]odality refers to both a semantic and grammatical notion. It can be expressed in a number of ways, but the most common ways are through verb inflection (*mood*) and through *modal* verbs or particles” (Nordström 2010, 16; author's italics). Modality focuses on the speaker and her perception of the world: “Modality differs from tense and aspect in that it does not refer directly to any characteristic of the event, but simply to the status of the proposition” (Palmer 2001, 1). Bybee et al. (1994, 176) observe that one overarching definition for all meanings and usages of Modality is impossible, or at least can only be found through diachronic study, but although they do present such a study, they do not present a new definition. Likewise, this chapter attempts to investigate the systematic characteristics of Modality in the SAILS sample rather than arriving at a single encompassing definition.

There is continuing discord about the inner workings of Modality. Narrog (2005, 189) refers to it as the “‘dustbin’ of grammatical categories” where many fuzzy categories tend to end up, but without a sufficiently coherent internal definition that makes it a valid category. Several attempts have been made to arrive at such a definition, e.g. by Palmer (1986, 2001) who divides it into two domains (Propositional and Event Modality), and Bybee & Fleischman (1995)¹ who distinguish three domains (Speaker-oriented, Agent-oriented, and Epistemic Modality). These divisions partly overlap, e.g. Palmer’s Propositional Modality includes Bybee & Fleischman’s Epistemic Modality, although he also includes Evidentiality. The present study does not explicitly follow one and disregards the other approach, but will be largely based on Bybee & Fleischman (1995). Their terminology will now be outlined and contrasted with that of Palmer (1986, 2001).

Agent-oriented Modality refers to conditions put on an agent whereas Speaker-oriented Modality has the agent putting conditions on an addressee. Thus, for Palmer, Modalities like IMPERATIVE and PROHIBITIVE that are external to the addressee (Deontic) belong to Event Modality. Modalities internal to the agent (Dynamic), like ability or willingness, also belong to Event Modality. Bybee & Fleischman subsume Deontic and Dynamic under Agent-oriented Modality as conditions that apply to the agent, whether imposed by him/ her or others, and establish another category, Speaker-oriented, for every Modality that has directive function. This last category includes speech-acts such as IMPERATIVE, JUSSIVE,

¹ Based on Bybee (1995).

HORTATIVE, OR PROHIBITIVE.² In general, the questionnaire tries to avoid broad divisions such as Epistemic, Deontic, etc., and instead asks for very specific markers, e.g. one will find questions 2.11 and 2.12 (“Is certainty marked morpho-syntactically?”/ “Is dubitative marked morpho-syntactically?”) instead of asking for Epistemic Modality. The specific definitions can be found in the relevant sections below. One point I disagree on with Palmer (2001) about is that I do not consider Evidentiality part of Modality, but as a separate category (see chapter 6).

Both Palmer’s and Bybee & Fleischman’s classifications have many things in common and indeed complement each other. As the present study is epistemological rather than theoretical, I prefer definitions that sprung from similar approaches, such as the approach by Bybee & Fleischman (1995), i.e. modeled on a survey by Bybee et al. (1994). Their language sample includes 76 languages (ibid. 31) (with only a small number of SAILS).³ Apparently, Palmer (2001) did not conduct a study of similar scale, but nevertheless bases his theory on sufficient data to be helpful in the present discussions.

In 2005, Narrog presented a definition of Modality that was supposed to finally give that term cross-linguistic validity. He identifies three approaches to Modality: i) speaker’s attitude (subjectivity), ii) factuality, and iii) everything non-propositional. Concerning a study of German, English, and Japanese he arrives at the conclusion that “the most reasonable cross-linguistic oriented definition of modality is one through the concept of factuality” rather than speaker’s attitude (ibid. 189). While essentially agreeing with his argumentation, for the sake of this study I feel it is too narrow. In the future, it may be useful, however, to carry out a study based on Narrog’s definition to see whether it is valid across SAILS.

The Modalities in the present study were chosen and others excluded for different reasons. A complete study of all possible Modalities is tempting but has to wait. It is attempted here to present a mixture of the new and appealing with the more traditional and well researched Modalities while taking into account the temporal and spatial limitations of a PhD project. On the one hand, some Modalities have a continuous history of research, recently received an increasing amount of attention, and are thus ideal for comparison, such as REALIS/ IRREALIS (cf. Bybee et al. 1994, Givón 1994, Palmer 1986 and 2001, Elliott 2000, Narrog 2005, Nordström 2010) or COMMAND types (cf. Aikhenvald 2010). On the other hand, the appeal of e.g. DESIDERATIVE and FRUSTRATIVE lies in their underrepresentation in the literature.⁴ Other, well known, Modality categories such as indicative, subjunctive, optative, necessity, ability, obligation etc. are not taken into account here. This presents difficulties for the final results, because this study investigates only a part of the Modal systems of the SAILS, in contrast to the Tense and Aspect investigations in the chapters above, which are

² Nordström (2010) prefers the term Speech-act Modality over Speaker-oriented.

³ Island Carib, Cocama, Chacobo, Jivaro, Tucano, as well as languages from the Macro-Chibchan and Gê-Pano-Carib groups (Bybee et al. 1994, appendix A).

⁴ Khanina (2008) is the one exception for DESIDERATIVE.

more or less complete studies of absolute Tense and viewpoint Aspect. Any quantitative result must therefore be regarded with double caution: First, because of the limitations of the data, and second, because of the limitations of the chosen Modality categories.

The following sections investigate in detail the systematic distribution of *REALIS/IRREALIS* (5.2), *CERTAINTY/DUBITATIVE* (Epistemic Modality) (5.3), *INTENTIONAL* (5.4), *POTENTIAL* (5.5), *PURPOSIVE* (5.6), *FRUSTRATIVE* (5.7), and *DESIDERATIVE* (5.8). Section 5.9 analyzes the *COMMAND* types *IMPERATIVE*, *HORTATIVE*, *JUSSIVE*, *POLITE COMMAND*, *COLLECTIVE COMMAND*, and *PROHIBITIVE*. Sections 5.10 – 5.12 present an analysis of Modality marking in the sample according to typological (5.10), geographical (5.11), and genealogical (5.12) distributions. A short comment of the implications of the findings for the stability of Modality is given in 5.13. The results are summarized in section 5.14.

Table 5.1: Classification of the Modal categories studied here

Modal category present in the sample	Classification according to Bybee et al. (1994)	Classification according to Palmer (1986, 2011)
<i>REALIS/IRREALIS</i>	-	Mood
<i>INTENTIONAL</i>	Agent-oriented Modality	Event Modality
<i>POTENTIAL</i>	Epistemic Modality	Event Modality
<i>CERTAINTY</i>	Epistemic Modality	Propositional Modality (epistemic)
<i>DUBITATIVE</i>	Epistemic Modality	Propositional Modality (epistemic)
<i>FRUSTRATIVE</i>	-	-
<i>PURPOSIVE</i>	subordinate mood	Event Modality (subjunctive)
<i>DESIDERATIVE</i>	Agent-oriented Modality	Event Modality (deontic/ epistemic)
<i>COMMAND</i> types	Speaker-oriented Modality	Event Modality (deontic)

5.2 *REALIS/IRREALIS*

5.2.1 Definitions

REALIS and *IRREALIS* are perhaps the most controversial terms within Modality. Bybee et al. (1994, 238) even propose that they may not be valid cross-linguistically, as they vary too much between languages. Givón (1994, 322) argues against this and states that complexity does not necessarily point to absence of a common core meaning. On the contrary, a closer study of *IRREALIS* reveals a “considerable measure of coherence and commonality” (ibid.). I follow Givón in this, though keeping an eye on the variation.

Previously, the typological categories *REALIS* and *IRREALIS* were taken to be a distinction between the real and unreal, unsatisfactorily so, as they depict what a speaker perceives and/or proposes to be real or unreal, rather than being absolute statements about the world (Palmer 2001, 2). Palmer subsumes both the indicative-subjunctive and *realis-irrealis* distinctions under the supercategory *Realis-Irrealis* as a subcategory of Modality. In 1986, he defined the subcategories *realis* and *irrealis* in terms of the semantic notion of factuality,

but later revised this to the speech-act notion of assertion and non-assertion, because factuality, in his opinion, cannot account for all usages (2001). Nordström (2010, 46) cautions against a REALIS/ IRREALIS speech-act definition on the grounds that it is conceptually too broad to be valid anymore. Instead, she proposes that REALIS/ IRREALIS are members of Propositional Modality (and not speech-act or Event Modality). They also are to be kept apart from declarative/ subjunctive, which also belong to Propositional Modality, but are not to be confused with REALIS/ IRREALIS. Nordström (ibid.) does concede that REALIS/ IRREALIS and declarative/ subjunctive together make up a (Propositional Modality) category of Realis/ Irrealis. Narrog (2005, 184), in an attempt to cleanse the discussion of unfortunate terminology, subsumes REALIS and IRREALIS under the term “factual” which he deems to be the best definition of Modality as a valid category. Givón (1994, 269) claims that besides the cognitive function (subjective certainty), REALIS/ IRREALIS also contribute towards communication functions, i.e. pragmatics, and that any theory about REALIS/ IRREALIS has to take this into account. Givón also claims that instead of characterizing a binary opposition, REALIS/IRREALIS work along a continuum (cf. also Van Gijn & Gipper 2009). What has been said so far shows that the dilemma in defining REALIS/ IRREALIS is how to specify them in enough detail to demarcate them from other categories, but also to be broad enough to be universally valid (cf. De Haan 2012 for a detailed history of the notion of irrealis). Studies on REALIS/ IRREALIS have been carried out for individual languages in many parts of the world, but never in a comprehensive way for SAILS.⁵ It is therefore attempted here to work with a definition that enables a comparison with other studies. The definition will be outlined in the following section.

In this thesis, the focus is on semantic categories. The main semantic function of REALIS/IRREALIS marking is to denote a proposition the speaker deems \pm factual. I also accept the possibility of REALIS/ IRREALIS to interact with and indeed take over speech-act functions: I follow Elliott (2000, 67) who gives the following semantic definition: “Prototypically realis is used in clauses where there is perceived certainty of the factual reality of an event’s taking place, while irrealis is used to identify that an event is perceived to exist only in an imagined or non-real world”. This, plus the following points characterize REALIS/ IRREALIS in this study:

1. A REALIS OR IRREALIS marker must express the attitude of the speaker towards (non-) factuality. Languages may vary where they put the cut-off point between factual and non-factual (i.e. what is REALIS in one language may be IRREALIS in another).
2. A REALIS OR IRREALIS marker must not mark a certain syntactic construction. A subordinator is not a REALIS/IRREALIS marker, but it must be taken into account that there may well be cumulative morphemes.⁶

⁵ See Elliott (2000) for Australian, Austronesian, and Papuan, and Mithun (2001) and Chafe (1995) for North American languages.

⁶ See Nordström (2010, 25) for this problem.

Obligatoriness of the category is not necessary. The ideal *REALIS* or *IRREALIS* distinction in a language is marked obligatorily,⁷ because it is then more easily detected. However, because the sources rarely discuss obligatoriness this cannot be taken into account with the present data (except where specified in individual cases).

Languages with a *REALIS*/*IRREALIS* distinction often mark only one of these overtly, the other being the default unmarked one. A noteworthy exception is Nanti (Arawakan), which obligatorily codes for reality status. *REALIS* is marked by suffixation and *IRREALIS* by circumfixation; both categories have two different forms whose selection is based on the class of the verb they attach to (Michael 2008, 250, 254): “realis marking is associated with positive polarity indicative of non-future temporal reference [...]; while irrealis marking is associated with future temporal reference [...], with negative polarity [...], and with counterfactual modality”.

5.2.2 Distribution in the sample

Only three languages in the sample mark *REALIS* morpho-syntactically, but 28 mark *IRREALIS*. All three languages that mark *REALIS* also mark *IRREALIS*. The form of marking ranges from suffixation over cliticization to particles; there are two prefixes for *IRREALIS* (in Panare and Karitiana).

The great imbalance of *REALIS*/*IRREALIS* (3:63 vs. 28:63) could be explained by assuming that *REALIS* usually is the default category, which does not receive overt marking. In Mapuche and Mosestén *REALIS* is indeed identified as the absence of *IRREALIS* markers (Smeets 2008, 235; Sakel 2004, 342). The following discussion first concentrates on the three languages marking *REALIS*, then on *IRREALIS*, taking their various associated functions into account.

Table 5.2: Distribution of *REALIS*/*IRREALIS* marking in the SAILS

<i>REALIS</i>	Wichí, Tehuelche, Wari'
<i>IRREALIS</i>	Panare, Wichí, Tehuelche, Wari', Huallaga Quechua, Timbira, Urarina, Karitiana, Paresi, Bororo, Baure, Jarawara, Awa Pit, Apurinã, Mapuche, Eméillon, Imbabura Quechua, Cubeo, Mamaindê, Warao, Kwaza, Kamaiurá, Mosestén, Tsafiki, Yurakaré, Mekens, Tiriyo, Cuzco Quechua

REALIS and IRREALIS marking

The three languages marking both *REALIS* and *IRREALIS* Wari', Tehuelche, Wichí, are discussed in the following. These languages also are the only ones marking *REALIS* and Wari' and Tehuelche incidentally also do not mark any other Modality (Wichí additionally marks *IMPERATIVE*, *PROHIBITIVE*, and *HORTATIVE*; see section 5.3).

The *REALIS*/*IRREALIS* system in Tehuelche is independent of Tense marking, which is only partially the case for Wari' and not at all in Wichí. Both the Tehuelche *REALIS* suffix *-k'* and

⁷ See e.g. Elliott (2000) who only includes obligatory *REALIS*/*IRREALIS* marking.

the IRREALIS suffix *-m* are compatible with Tense marking, but do not themselves express Tense (Fernández Garay 1998, 274). In contrast, Wichí has a REALIS/ IRREALIS distinction that occurs only in the FUTURE. The two cumulative morphemes *hila* and *hina*⁸ express REALIS FUTURE and IRREALIS FUTURE, respectively. The clitic *hila* is used for actions which presumably will be realized, whereas with *hina* the probability is quite low. The following examples differ only in their usage of the clitics; in (5.1a), the speaker transmits great certainty that he is going to buy clothes, whereas in (5.1b), he is less certain. The latter can be interpreted as a promise or an announcement, but with a low probability (Terraza 2009, 164, 165). The difference between the two clitics can also be viewed as one where the speaker insists on the probability of an action he will carry out (*hina*) vs. focusing on the fact that the action will occur (*hila*) (Terraza 2009, 165). It is therefore debatable whether the latter form really carries Modal value, or if the difference between *hina* and *hila* is rather one between a Modal FUTURE and a FUTURE without Modal value.

(5.1) Wichí (Matacoan; Terraza 2009, 165)

- (a) *n-k'ox-u-hila* *to-wey-is* *n-k'ila*
 1-buy-APPL-FUT.REA POSS.IND-clothes-PL 1POSS-older.brother
 'I will buy clothes for my older brother.' (orig. 'Voy a comprar ropa para mi hermano mayor.')
- (b) *n-k'ox-u-hma* *to-wey-is* *n-k'ila*
 1-buy-APPL- FUT.IRR POSS.IND-clothes-PL 1POSS-older.brother
 'I will buy clothes for my older brother.' (orig. 'Voy a comprar ropa para mi hermano mayor.')

In Tehuelche, the choice of REALIS or IRREALIS depends on what the speaker deems factual or non-factual. The IRREALIS suffix *-m* used in interrogative and subordinate clauses, to express possibility and doubt, and when the speaker is not absolutely certain about the proposition. The REALIS suffix *-k'* is used for facts known by the speaker, or when he is absolutely sure. The choice of the Tehuelche REALIS/ IRREALIS is in part based on the knowledge of the speaker. When the speaker is certain or has knowledge about what he is going to utter in the proposition, he uses the REALIS suffix, but when he is uncertain he uses IRREALIS.

Wari' marks a factuality continuum ranging from REALIS on the one side to IRREALIS on the other. REALIS occurs in cumulative morphemes with Tense, either PAST/ PRESENT OR FUTURE (and number and person), but IRREALIS does not. According to which degree the speaker deems the event/ action factual or not, he chooses from the REALIS or IRREALIS side. In the middle of the continuum there is a REALIS FUTURE marker, which is interesting as FUTURE is a meaning in other languages usually associated with IRREALIS. In Wari', FUTURE REALIS can mean

⁸ These particles can occur discontinuously, i.e. additional verbal morphology can occur between the element *hi-* and either *-la* or *-ha*. Nevertheless, Terraza considers them to be single morphemes each (Terraza 2009, 165).

that the speaker is certain or uncertain, according to context (Everett & Kern 1997, 313). The REALIS/IRREALIS markers are the only Modality markers in Wari' as coded by the questionnaire and all other Modalities in Wari' are marked either non-morpho-syntactically or expressed by a combination of REALIS/IRREALIS marking with other means, such as intonation or context. The Wari' indicative can be expressed by the use of REALIS PAST/PRESENT forms, and INTENTIONAL can be expressed by the REALIS FUTURE (Everett & Kern 1997, 317f.). The verbalized desiderative clause can only take IRREALIS markers (ibid. 65). IRREALIS marking can be used to signal conditional clauses, even though an alternative marking with *mo* 'conditional' exists (ibid. 101). The particle *mo* and REALIS/IRREALIS marking can co-occur.

(5.2) Wari' (Chapakuran; Everett & Kern 1997, 317, 99, 101)

- (a) *cao' nana-in mijac 'oro wari'*
eat 3PL.REA.PST/PRS-3N pig COL person
'The people ate the pig.'
- (b) *mo xi cao' ca. Ma'am xir*
COND INFL:IRR eat 3SG.M full 3SG:IRR
'If he would eat, he would be full.'
- (c) *mo xira xi taxi' ca*
to:SG 3SG:IRR INFL:IRR know 3SG.M
'He would have gone if he had known.'

The investigation of REALIS/IRREALIS above also shows that these categories are highly related to Epistemic Modality and Tense and often quite difficult to distinguish from CERTAINTY/DUBITATIVE/FUTURE markers. It seems that in some cases the terminology depends on the preference of the author and it is observed that sometimes these terms are used almost synonymously. The present study attempts to make a distinction between them, but it should be borne in mind that this is not always possible due to limitations of data.

IRREALIS marking

The following section focuses on the different functions of IRREALIS marking in the sample. Languages vary according to how many functions IRREALIS markers can have. IRREALIS can be quite restricted, as e.g. in Apurinã, where the IRREALIS suffix *-ã* only marks a hypothetical proposition (Facundes 2000, 325). Another example is the IRREALIS particle *na* in Awa Pit (which Curnow 1997 glosses as 'counterfactual irrealis'), which is restricted to the apodosis of a past counterfactual statement (ibid. 277). IRREALIS can also occur in syntactically less restricted positions, as e.g. in Jarawara, where the IRREALIS occurs in conditional, but also in non-conditional clauses (Dixon 2004, 215):

(5.3) Jarawara (Arawan; Dixon 2004, 216, 215)

- (a) *Jobeto bija kihe-himana-ha,*
 name.M battery.F have-HYP.M-DEP
bija taa ne-na ama-ka
 battery.F give AUX-IRR.M EXTENT-DECL.M
 'If Jobeto had batteries, he would give batteries (to us).'
- (b) *ke-tehe-ne ama-ke jaha jaa*
 APPL-be.lubricated-IRR.F EXTENT-DECL.F oil PERI
 'It (some piece of machinery) needs to be lubricated with oil.'

Many IRREALIS markers in the sample have, in one way or another, FUTURE connotations. This can manifest itself in cumulative (cf. Wichí above) or ambiguous morphemes, or IRREALIS forms expressing FUTURE in languages with no separate FUTURE marking. This does not exclude the possibility that an IRREALIS marker can also be used with past time reference. Whereas events in the future are naturally not factual (yet), events in the past that did not come true are also not factual (and will never be). However, it is much more common for an IRREALIS marker to be used for FUTURE than PAST. The following paragraphs take a closer look at the distribution of IRREALIS marking and Tense/ time reference as well as other Modalities in the sample.

Timbira has a particle *ha*, which marks FUTURE, but also IRREALIS (Alves 2004, 67). Alves (2004) claims that this marker has Modal value besides marking FUTURE Tense, as can be seen in (5.4b) where *ha* occurs in a conditional clause. It is unclear whether *ha* is ambiguous between FUTURE and IRREALIS, or a cumulative morpheme.

(5.4) Timbira (Macro-Gêan; Alves 2004, 67)

- (a) *wa ha tẽ*
 1 FUT viajar
 'I am going to travel.' (orig. 'Eu vou viajar.')
- (b) *ke ha ku-tʃɜ nã mã ke ha ramã rɔp kura*
 3 IRR 3-morder SUB DSBJ 3 IRR ASPC cachorro matar
 'He would have killed it, if the dog would have bitten him.' (orig. 'Ele o mataria se o cachorro o mordesse.')

In (5.5) from Jarawara, a statement about a PAST event that did not come true is marked with IRREALIS; in Mosestén, the IRREALIS marker can have past time reference but can also apply to a FUTURE event (5.6); the BORORO IRREALIS suffix *-mëdĩ* can refer to any time reference (5.7)⁹; and in Karitiana, the IRREALIS prefix *i-* can occur both in clauses with FUTURE and NONFUTURE marking (5.8):

⁹ This suffix is not mentioned in Nonato (2008).

- (5.5) Jarawara (Arawan; Dixon 2004, 215)

mee ka-so-hene-ke [jama.kabani jaa]
 3N-SG.SBJ APPL-fall-IRR.F-DEC.F forest(F) PERI

'They (in an aeroplane) could have crashed into the forest, but did not.'

- (5.6) Mosetén (Mosetenan; Sakel 2004, 342, 343)

- (a)
- yäe-rä'*
- bojw-i*

1SG-IRR go.up-VM.M.SBJ

'I will go up.'

- (b)
- mäen'jä'-khan nāsh-tsa' aj mi'-chhe-in karij-tya-ki-wi'-ra'-in*

yesterday-IN FOC-FRUST yet 3M-SU-PL hard-VM-ANP.M.SBJ-M-IRR-PL

'The other day they should already have worked (but they did not).'

- (5.7) Bororo (Macro-Gêan; Crowell 1979, 92)

imedi maragodi-mëdi

man(3SG) work-IRR

'The man probably worked.' or 'The man probably is working.' or 'The man probably will work.'

- (5.8) Karitiana (Tupían; Everett 2006, 329, 255)

- (a)
- i-ator-i ñn bĩpãn*

IRR-take-FUT 1SG arrow-OBQ

'I will not take the arrow.'

- (b)
- ñn i-opĩ-t (ñn)*

1SG IRR-cut-NFUT 1SG

'I did not cut it.' or 'I did not cut.'

It could be argued that in languages that do not have FUTURE, but do have IRREALIS marking, the latter is rather a Tense marker. Apart from the fact that an IRREALIS can cover, but is not necessarily restricted to FUTURE reference, this also not supported by the fact that there are languages with both categories marked. In Karitiana, the FUTURE suffix and IRREALIS prefix can even co-occur (though arguably the IRREALIS contributes to the negative meaning rather than FUTURE) (5.8a). It is of course also possible that a FUTURE marker can express IRREALIS in languages with no separate IRREALIS marking, as in e.g. Apurinã (Facundes 2000, 515).

As was already demonstrated for Tehuelche, Wichí, and Wari', the choice of (non-) factuality marking can depend on the certainty of the speaker of the proposition. In Bororo, the IRREALIS marker itself already indicates uncertainty, but it can also be combined with the DUBITATIVE marker to express an even greater uncertainty (Crowell 1979, 93):

- (5.9) Bororo (Macro-Gêan; Crowell 1979, 93)

e-mago-ru-mēde

3PL-speak-DUB-IRR

'They may perhaps speak to him.'

Another frequent phenomenon in the sample is the ability for IRREALIS markers to cover/interact with Speaker-oriented Modality. For example, the Wari' IRREALIS clitic can be used for polite/mild commands. Wari' does not have any morpho-syntactic marking for IMPERATIVE, so the IRREALIS clitic can be used instead for polite IMPERATIVE instead. In Baure, the IRREALIS suffix *-ša* also refers to polite requests/ commands, or suggestions (Danielsen 2007, 347).

- (5.10) Wari' (Chapakuran; Everett & Kern 1997, 35)

'U'um 'o *xima-ocon* waram wia-xi'

dawn;dusk lead 2SG:IRR-3PL.M COL grandchild-1PL.INCL

'You should take our grandsons early in the morning.'

- (5.11) Baure (Arawakan; Danielsen 2007, 347)

pi=epitia-ša=ni

2SG=do.favour-IRR=1SG

'Please! (lit. do me a favour!)

In Karitiana, Baure, and Tiriyo the respective IRREALIS markers can co-occur with IMPERATIVE marking. The Karitiana IRREALIS prefix *i-* is obligatory in forming an IMPERATIVE from a transitive verb, co-occurring with the IMPERATIVE suffix *-a* (Everett 2006, 326). The Tiriyo IRREALIS particle *mo*, when combined with IMPERATIVE marking, expresses a command that has to be carried out later in contrast to right now (Meira 1999, 454).¹⁰

- (5.12) Karitiana (Tupían; Everett 2006, 379)

(ān) i-okēn-a

2SG IRR-cut-IMP

'Cut it!'

- (5.13) Tiriyo (Cariban; Meira 1999, 455)

ëwě_pā_mo apěh-ke

later_REP_IRR get-IMP

'Get it later (not now).'

¹⁰ In Baure, the co-occurrence of IRREALIS and IMPERATIVE marking does not visibly change the command meaning

Two languages in the sample show peculiarities not found elsewhere: Masetén and Emérillon. Masetén has a basic REALIS/IRREALIS distinction, but only the IRREALIS is marked, by the suffix *-ra*. It occurs with FUTURE reference, in rhetorical questions, and negative clauses, but also marks a less-typical feature not usually associated with non-factuality: change in viewpoint (Sakel 2004, 342ff.).

The IRREALIS particle *itje* ~ *ite* in Emérillon is mostly used in subordinate clauses, but when it occurs in independent clauses it has a meaning otherwise associated with FRUSTRATIVE (Rose 2003, 438):

(5.14) Emérillon (Tupían, Guaraní; Rose 2003, 438)

o-pokadz-a-itje, wai o-koal-a-öwã
3-shout-a-IRR NEG 3-find-a-?

'He shouts in vain, he does not find her.' (orig. 'Il a beau crier il ne la trouve pas.')

A remark on negation is called for here. The present study focuses on positive clauses, but it should be noted that many of the IRREALIS markers occur in negated clauses (but are not themselves negators).

The function of IRREALIS as conditional/counterfactual marker may lead to an even greater number of IRREALIS markers in the sample than presently identified; they are hidden in grammars as contrary-to-fact markers. On the other hand, if one were to exclude conditional markers from the IRREALIS definition, the number of IRREALIS markers in the sample would decrease greatly.

The previous paragraphs have attempted to give an overview of REALIS/IRREALIS marking in the sample. There are IRREALIS markers interacting with Tense, Epistemic Modality, and Speaker-oriented Modality. This is evidence in favour of Palmer's (2001) definition; Nordström's (2010) definition is too narrow to account for all these functions. It is not the purpose of this study to finally answer the question whether there is a universal REALIS/IRREALIS category, but it is hoped that the present data contributes to the discussion. There certainly is no unified picture of REALIS/IRREALIS marking in the SAILS sample according to the definitions above.

5.3 CERTAINTY/ DUBITATIVE (EPISTEMIC MODALITY)

5.3.1 Definitions

Speakers can choose to state in a proposition whether or not they think something is likely, or whether they believe that an event may come true or not. This value is called Epistemic Modality and reflects the opinion of the speaker: "Epistemic Modality applies to assertions and indicates the extent to which the speaker is committed to the truth of the proposition" (Bybee et al. 1994, 179). Epistemic Modality markers exist in a continuum between certainty and doubt. The speaker uses a marker for certainty when he is sure of the truth of the proposition, and doubt when he is not sure. The degree of being (not) sure can vary from very certain over less certain to not really certain and completely uncertain. For example,

Mosetén has a quite complex system of expressing certainty, uncertainty, and shades in between. Two examples for high certainty and moderate or low certainty are:

(5.15) Mosetén (Mosetenan; Sakel 2004, 347, 353)

- (a) **chhata'** yäe rai's-e-' jaem'-ñae'-tye-'
 CERT 1SG want-VM-3F.OBJ good-VM-APPL-3F.OBJ
 'I truly want to make it.'
- (b) **watyeke** jao-tya-k-wa-'jo-i
 DUB smoke-VM-ANP-PROG-PRI-M.SBJ
 'Maybe he is smoking (fish).'

Not included here are markers that occur exclusively in interrogative clauses, because of the probability that they are actually interrogative markers. For example, Dâw has a particle *?e* labelled 'DUBITATIVE' that is restricted to interrogative clauses:

(5.16) Dâw (Nadahup; Martins 2004, 483)

- hũ?* *?a* *?e*
 who be DUB
 'Who would/ could that be?' (orig. 'Quem sera esse?')

CERTAINTY in this thesis refers to the commitment to the truth of the proposition by the speaker; DUBITATIVE refers to a marker that includes any kind of doubt. A DUBITATIVE can therefore be a marker that marks an event with a high degree of certainty, but that still has some elements of doubt in it. It is typically translated into English with 'may(be)', 'I guess', 'I think', 'it seems', etc. DUBITATIVE therefore allows for a variety of markers along the certainty-continuum, whereas CERTAINTY is very restricted. This is easy to explain; there is a natural state of complete certainty (which does not allow for inner subdivisions, either one is certain or not), but not of complete uncertainty. The continuum can thus be regarded as open-ended on the uncertainty side.

CERTAINTY and DUBITATIVE values are related to direct and indirect Evidentiality respectively, as an event the speaker did not witness can be interpreted as being doubtful, and an event that was directly witnessed as certain. Markers may not always be clear whether they are Evidentials with secondary Epistemic meaning, or true Epistemic Modals. There is a great potential for Epistemic Modality and Evidentiality to overlap, e.g. "an indirect evidential, which indicates that the speaker has only indirect knowledge concerning the proposition being asserted, implies that the speaker is not totally committed to the truth of that proposition and thus implies an epistemic value" (Bybee et al. 1994, 180). One should bear in mind that a marker that primarily signals certainty or uncertainty even if based on some sort of evidence is still coded as Modality marker here, whereas a marker that signals source of information primarily is an Evidential marker,

regardless of whichever conclusions referring to CERTAINTY status can be drawn (for more about Evidentiality see chapter 6).

5.3.2 Distribution in the sample

In the present sample of 63 languages 35 mark Epistemic Modality, 20 of which mark CERTAINTY and 29 mark DUBITATIVE. Their distribution is presented in table 5.3. This section takes a closer look at Epistemic Modality systems and degrees of complexity, Epistemic Modality and its relationship to Tense, and origins of individual Epistemic forms; geographical/ genealogical distributions are discussed in section 5.10.2.

Table 5.3: Distribution of Epistemic Modality marking in the SAILS

Marking	Number	Languages in the sample
CERTAINTY only	6	Baure, Apurinã, Aymara, Rikbaktsa, Dâw, Mamaindê
DUBITATIVE only	15	Paresi, Tariana, Kaingang, Bororo, Hup, Matses, Imbabura Quechua, Cavineña, Desano, Cubeo, Sateré-Mawé, Karo, Emérillon, Kamaiurá, Urarina
Both	14	Mapuche, Yanasha', Tsafiki, Hixkaryana, Tiriyo', Mose'tén, Yaminahua, Cuzco Quechua, Mekens, Cocama-Cocamilla, Nheengatú, Puinave, Movima, Yurakaré
None	28	Jarawara, Awa Pit, Miraña, Panare, Wari', Ika, Chimila, Embera, Tehuelche, Mocoví, Pilagá, Aguaruna, Timbira, Wichí, Sabanê, Nasa Yuwe, Shipibo-Konibo, Huallaga Quechua, Karitiana, Tapiete, Itonama, Kanoê, Trumai, Munich, Leko, Warao, Kwaza, Yanam

Epistemic Modality systems

Languages vary according to whether they have more than one morpheme for the same Epistemic Modality. Simple systems have one marker only, as e.g. in Baure or Tariana, whereas there can be two or more markers within the same language for roughly the same meaning (e.g. Tiriyo'). Most of the languages show a one-to-one correspondence of morpheme and meaning, but in e.g. Cubeo and Hixkaryana Epistemic Modality is marked in cumulative morphemes with Tense.

Simple systems are found where one language marks either CERTAINTY or DUBITATIVE with a single marker, but more interesting than single-marker languages are those that have more than one marker, because of their internal semantic distribution. What prompts the existence of several markers for the same meaning? Certainty and doubt exist in a continuum with many shades between them and it is to be expected that in languages with a range of Epistemic markers these cover different areas. Given the fact that certainty and doubt are semantic fields that relate to Tense and Evidentiality one may assume that these factors also play a role in the distribution. This section investigates complex (i.e. more than one marker per category) Epistemic Modality systems and their entanglements with other categories.

The language with the most complex Epistemic Modality system in the sample is Mose'tén. Mose'tén has a whole range of Epistemic Modality markers: Three markers express CERTAINTY, and four markers express DUBITATIVE. Their internal distribution is not so clear,

however. The difference between the two high CERTAINTY markers particle *chhata'* and the particle *ika'/ öka'* is at first not obvious. Both express a high degree of certainty to the point where the speaker is sure about what he says. None of the markers seems restricted to a certain Tense or clause type (the particle *chhata'* often occurs in FUTURE contexts but is not restricted to them). The particles *ika'/ öka'* can occur in rhetorical questions, a remnant of their origin as the pronominal forms *i* and *ö* plus the rhetorical question marker *-ka* (Sakel 2004, 347), but also in declarative clauses, and, to confuse matters, they can also express less than complete CERTAINTY (ibid. 348). They often occur in contexts with names and the speaker, not remembering the name(s), asks himself this rhetorical question:

(5.17) Mosetén (Mosetenan; Sakel 2004, 348)

tyi-ra' ö-ka' mö' tí'-i'?
 person-IRR CERT 3F.SG name-VM-F.S
 'What was her name?'

It is likely that this rhetorical construction is the origin of *ika'/ öka'*, which have then developed into UNCERTAINTY markers and finally CERTAINTY markers. This special path of grammaticalization still has to be explored. Apparently, these particles and *chhata'* do not co-occur. A more thorough investigation is necessary to see whether they are in complementary distribution.

The DUBITATIVE markers =*tyi'*, *jäedyäk* and *watyeke* all refer to a moderate to low level of certainty. They do not seem to be restricted to a certain Tense either. Rather than occupying different slots in the certainty continuum, they differ in indicating why the speaker is uncertain. Using *watyeke*, the speaker has no grounds on which to base his judgement, whereas with *jäedyäk* he may have some Evidential input, but not necessarily. Using the clitic =*tyi'*, the speaker states that he has no personal experience on which to base his judgement (cf. Sakel 2004, 350-351), which is quite a strong factor for being uncertain. This clitic can even devalue *chhata'*; when both co-occur, the meaning changes to a possibility (ibid. 347). This clitic also occurs in the clitic =*dya'tyi'*, which expresses uncertainty based on unclear evidence (consists of the bound morpheme *dya'* and the DUBITATIVE clitic =*tyi'*) (ibid. 353). Table 5.4 below illustrates the Epistemic markers in Mosetén and ranks them from high certainty to low certainty.

To summarize, the Epistemic Modality markers in Mosetén cover roughly the same area of moderate certainty, but differ in what evidence there is, or rather is not, for the speaker's judgement. This is closely related to the realm of Evidentiality, but they are not Evidential markers as such, because they do not specify how the information was acquired (cf. chapter 6).

Mosetén is an example of a two-sided Epistemic continuum, in which again there is a distinction made according to information source. In Tiriyó, there seems to be a similar system. Occupying the slot on the side of complete CERTAINTY is the suffix *-e*, while the suffix *-ne/ -në* and a group of particles occupy the DUBITATIVE area (*tahke(ne)*, *tahkarë*, and *tahkara*).

Not much is known about possible differences in meaning within the group of particles; they all imply uncertainty (Meira 1999, 459), but the forms *tahkarë* and *tahkara* both code “that the speaker is not sure, but that something in the situation is leading him/ her to the conclusion stated in the sentence” (ibid. 461) and therefore the markers code a deductive process. The particle *tahke(ne)* indicates a stronger doubt than the suffix (ibid. 460). They can co-occur, although it is not clear whether this increases the amount of doubt.

(5.18) Tiriyo (Cariban; Meira 1999, 312, 314)

- (a) *n-eh-ta-n* *konopo*
 3AO-come-FUT.IPFV-DUB rain
 ‘The rain will come.’
- (b) *wi-h-ta-n* *tahken*
 1S_A-go-FUT.IPFV-DUB maybe
 ‘Maybe I will go.’

A language where two markers clearly occupy different slots in the Epistemic continuum without referring to the source of the judgement, but co-occurring with Evidential markers, is Desano. It has two DUBITATIVE suffixes: *-sa* and *-sia*. The latter indicates less doubt than the former and occurs with VISUAL Evidential marking, whereas the former occurs with the ASSUMED Evidential (Miller 1999, 82).

(5.19) Desano (Tucanoan; Miller 1999, 82)

- (a) *wi?i-ge* *ári-sa-kū-bō*
 house-LOC be-DUB-ASM-3F.SG
 ‘She may be in the house; I don’t know for sure.’
- (b) *wi?i-ge* *ári-sia-bō*
 house-LOC be-DUB-3F.SG
 ‘She seems to be in the house.’

A last example for an Epistemic Modality system where the markers relate to Evidentiality is Dâw. It has two CERTAINTY markers: the suffix *-ih* ‘verification’ and the particle *tūn*. The latter occurs in affirmative clauses¹¹ and always with the verb *páh* ‘know’ (Martins 2004, 483). The suffix seems to be restricted to contexts where the speaker wants to emphasize his/ her certainty of the proposition. No example where both co-occur could be found.

¹¹ In contrast to the DUBITATIVE particle *?e* which is restricted to interrogative clauses (see 5.6.1) (Martins 2004, 483).

(5.20) Dâw (Nadahup; Martins 2004, 477, 483)

- (a) *cém* *ʔāh* *xu-ih*
 yesterday 1SG go.down-CERT
 'It is the truth! I came down the river yesterday.' (orig. 'È verdade mesmo! Eu vim rio abaixo ontem.')
- (b) *ʔāh páh* *tūn/* *tih* *wùd*
 1SG know CERT 3SG arrive
 'I think that he arrived already.' (orig. 'Eu acho que ele já chegou.')

Table 5.4: Epistemic Modality markers in Mosestén ranked from high to low certainty (adapted from Sakel 2004)

Marker	Form	Function	Epistemic Modality
<i>chhata'</i>	Particle	High degree of certainty	CERTAINTY
<i>ika/ óka'</i>	Particles (M/F)	Certainty	CERTAINTY
= <i>yai</i>	Clitic	Complete certainty	CERTAINTY
= <i>tyi'</i>	Clitic	Moderate certainty, possibility	DUBITATIVE
<i>jäedyäk</i>	Particle	Moderate to low certainty	DUBITATIVE
<i>watyeke</i>	Particle	Moderate to low certainty	DUBITATIVE
= <i>dya'tyi'</i>	Clitic	Uncertainty based on unclear evidence	DUBITATIVE

5.4 INTENTIONAL

5.4.1 Definition

INTENTIONAL is part of Agent-oriented Modality and refers to the plan or intention to carry out an action. Typically, this involves future time reference, as a planned action is one not yet carried out, but the difference between FUTURE and INTENTIONAL is that there is a will behind INTENTIONAL (cf. Agent-oriented Modality as to Bybee et al. 1994, 256). It is difficult to distinguish between FUTURE and INTENTION. Ideally, a FUTURE marker refers primarily to Tense (i.e. stating a fact without Modal values), and an INTENTIONAL marker to the intention of a speaker to carry out an action, but because that action is necessarily located at a time subsequent to the moment of the utterance, it includes FUTURE reference. INTENTIONAL typically occurs with the first person ('I want' constructions) as well as with verbs whose action is controllable, e.g. the verb 'to eat' is much more likely to be marked with INTENTIONAL than 'to faint'. This verb is much less controllable and therefore much less likely to be planned (depending on the circumstances). This is not the case with FUTURE marking, which is perfectly natural with all persons and verbs. It is expected that INTENTION and FUTURE are heavily intertwined. They may be marked by the same marker (cumulative or ambiguous morpheme), or one meaning is secondary to another. Which applies to which, has to be established for each case individually.

5.4.2 Distribution in the sample

Out of 60 languages in the sample, 13 mark INTENTIONAL morpho-syntactically (Baure, Tariana, Jarawara, Ika, Tiriyó, Aguaruna, Nasa Yuwe, Matses, Mamaindê, Itonama, Movima,

Yurakaré, and Warao).¹² The markers take the forms of suffixes, enclitics, particles, and, in one case, a prefix (Itonama). In most cases, there is a single INTENTIONAL marker per language. Baure is the only language exhibiting two INTENTIONAL markers: the particle *kač* and the suffix *-pa*. They can be combined or occur on their own without a change in meaning. The preverbal particle *kač* derives from the verb *kač* ‘go’. The suffix *-pa* is homophonous with the particle *pa* which occurs in IMPERATIVE constructions, and it is likely that they are related. Danielsen (2007, 292) even glosses the particle as “intentional imperative”. Baure does not have morpho-syntactic FUTURE marking and both *kač* and *-pa* can be used to express FUTURE time reference (ibid. 277). Baure does have an intricate system of marking IMPERATIVE (see section 5.3), in which the particle *pa* functions as emphazier.

(5.21) Baure (Arawakan; Danielsen 2007, 278, 292)

- (a) *ver kač eto=vi eto-a-pa=vi*
 PFV INT finish=1PL finish-LK-INT=1PL
 ‘We will already finish, we will finish.’
- (b) *pa nti’ nik=ro*
 GO 1SG 1SG.eat=3SG.M
 ‘I (emphatic) will eat it!’

Baure is not the only language where the INTENTIONAL marker may function as a command: the ambiguous INTENTIONAL/ DESIDERATIVE suffix *-bone* in Jarawara can be used for a mild IMPERATIVE and the Matses INTENTIONAL suffix *-nu* occurs in hortatory contexts. INTENTIONAL *-nu* almost always signals an action that is about to begin, with the initial steps already taken. The Ika INTENTIONAL marker can also signal an immediate FUTURE (Frank 1990, 60), as can the Tariana INTENTIONAL clitic *=kasu* (Aikhenvald 2003, 383).

(5.22) Matses (Panoan; Fleck 2003, 435, 437)

- (a) *pe-nu*
 eat-INT
 ‘I’m gonna eat now.’ (already putting the spoon in the soup)
- (b) *nid-nu nid-∅*
 go-INT go-IMP
 ‘Let’s go!’ (lit. ‘I’m going; go!’)

¹² The suffix *-gri* in Imbabura Quechua, which Cole (1982, 150) analyzes as ingressive Aspect, may also be an INTENTIONAL marker. Muysken (1977, 107) claims that *-gri* ‘going to’ has developed from the construction *V+kri-* ‘go’.

(5.23) Tariana (Arawakan; Aikhenvald 2003, 383)

dhinuru *di-thuka-kasu* *di-a*
 3SG.NF+neck 3SG.NF-break-INT 3SG.NF-go

'His neck is about to break.' (Said of a drunken man, lying on a bench with his head dangling.)

Tariana (5.23 above) and Yurakaré exhibit INTENTIONAL markers that can co-occur with non-control verbs. In Yurakaré, the INTENTIONAL suffix *-ni* occurs with a verb usually not associated with control over the action: 'to vomit'. As a result of nausea, vomiting is a body process out of control of the person concerned, although this point of view may be culture-specific.

(5.24) Yurakaré (unclassified; Van Gijn 2006, 197)

li-ti-ja-n-tta-ma *lumulu* *aye-ni*
 DEL-1SG-3SG-IDR.OBJ-put-IMP.SG warm vomit-INT:1SG.SBJ

'Put on some warm water for me, I am about to vomit.'

Unsurprisingly, all INTENTIONAL markers occur in clauses with FUTURE contexts (also in FUTURE in the PAST); indeed, the Movima particle *loy* is not allowed with PAST reference (Haude 2006, 527). The INTENTIONAL marker can co-exist with a FUTURE marker, or, if a language does not have FUTURE marking, it can take over that function (e.g. in Baure).

INTENTIONAL is also very close to DESIDERATIVE, which codes a wish or desire, but an action that is planned or intended does not necessarily involve the desire to do it. It is possible that DESIDERATIVE and INTENTIONAL are marked by the same suffix, like *-bone* in Jarawara:

(5.25) Jarawara (Arawan; Dixon 2004, 186, 185)

(a) *amo* *o-mata-ra-habone* *o-ke*
 sleep 1SG.SBJ-SHORT.TIME-NEG-INT.F 1SG-DECL.F
 'I don't intend to sleep for a while yet.'

(b) *o-ka-tima-ma-mata-habone*
 1SG.SBJ-in.motion-UPSTREAM-BACK-SHORT.TIME-INT.F
o-ke *mata*
 1SG-DECL.F SHORT.TIME
 'I want to go back upstream for a short time (to go back to my village to arrange things, before setting off on the journey to the Sorowahá).'

INTENTIONAL is closely related to commands (IMPERATIVE, HORTATIVE), FUTURE (immediate or general), and DESIDERATIVE. When the INTENTIONAL marker is used in IMPERATIVE contexts, it may signal a milder command in comparison to the usual IMPERATIVE constructions.

INTENTIONAL markers can serve to express FUTURE in languages that do not have a distinct FUTURE marking, but they can also be ambiguous with FUTURE.

These results are in line with research on non-SAILS. Bybee et al. (1994) cite Renck (1975, 228) for Yagaria, a language of Papua New Guinea, that the “Intentional Future form, [...] is also used to express imperative, hortative, and intention”. According to Bybee et al. (1994, 240-266), INTENTIONAL is a step before FUTURE on the grammaticalization path, regardless of what meaning came before intention.

For three languages in the sample there are origins given for the markers: Baure, Aguaruna, and Mamaindê. They will be discussed now in the light of possible grammaticalization paths postulated in Bybee et al. (1994).

The Aguaruna INTENTIONAL markers (suffix *-tatus* for 3rd person and suffix *-tasa* for non-3rd person) probably derive from the DESIDERATIVE suffix *-tatu ~ tata* plus the subordinator *-sa* (Overall 2007, 382). Bybee et al. (1994, 256) postulate the following grammaticalization path:

desire > (willingness) > intention > future

They do not include willingness as a full stage in between but rather as a nuance. Thus, the Aguaruna INTENTIONAL, like the Baure INTENTIONAL, may very well be on its way to grammaticalize into a FUTURE marker (given that those languages survive long enough to develop naturally). In Bybee et al.'s (1994, 253) sample, FUTURE markers most commonly derive from movement verbs such as ‘come’, or ‘go’. The fact that the Baure INTENTIONAL particle *kač* derives from the verb ‘go’ shows that it is possibly on the way to becoming a FUTURE marker.¹³ The Mamaindê INTENTIONAL suffix *-ten* seems to confirm the grammaticalization path above. According to Eberhard (2009), the suffix can be used as an alternative FUTURE marker without any intentional meaning left. It cannot co-occur with the FUTURE suffixes in either its FUTURE or INTENTIONAL carnation (Eberhard 2009, 420f.).

5.5 POTENTIAL

5.5.1 Definition

POTENTIAL is part of Epistemic Modality, in that it can depend on the speaker’s judgment whether an event is or was possible: “[p]ossibility indicates that the proposition may possibly be true” (Bybee et al. 1994, 179). The speaker himself may not be sure whether an event is true, because he has insufficient information. The event referred to then gets marked with a POTENTIAL marker; the English translation is usually the Modal verb ‘may’.

The POTENTIAL is expected to turn up in clauses with FUTURE marking, but also in clauses that refer to PAST events that may have, but did not, come true. Because it is conceptually close to being uncertain, it is also expected to find a relation to DUBITATIVE. Typical Modal environments would be IRREALIS, conditional, negative, and interrogative clauses.

¹³ Similarly, the Itonama FUTURE auxiliary developed from the verb *mama’na* ‘go’ (Crevels p.c.).

5.5.2 Distribution in the sample

Out of the 63 languages in the sample 17 mark POTENTIAL morpho-syntactically (Baure, Awa Pit, Nasa Yuwe, Matses, Yaminahua, Cavineña, Desano, Sateré-Mawé, Leko, Mekens, Yanam, Movima, Yurakaré, Kwaza, Kamaiurá, Puinave, Warao). The markers take the form of either particles or suffixes, with one noticeable exception: the POTENTIAL circumfix in Cavineña, which refers to FUTURE events and is “often used in questions to express soft and polite requests” (Guillaume 2008, 179). It is the only circumfix in the whole TAME dataset.

The languages do not vary much in how many POTENTIAL markers they have. With the exception of Matses which has as many as four different suffixes most have only one. Nasa Yuwe is the only language with a paradigm of POTENTIAL and person/ number cumulative morphemes. The following discussion focuses on the meanings of language specific POTENTIAL markers and in which environments they occur.

Unsurprisingly, the POTENTIAL can occur in clauses that are marked for FUTURE or have general future reference, though in some languages they are applicable only in future contexts, whereas in others the past is an acceptable context as well. The four POTENTIAL suffixes in Matses all have future reference. The suffixes *-mane*, *-nunda*, and *-panondac* differ only in which argument they take. All of them differ semantically from the last suffix *-nushe* in that they signal “that there is a possibility that the speaker will accidentally or unwittingly perform or cause an action” (Fleck 2003, 439) whereas *-nushe* additionally refers to a “future irrealis or counterfactual notion that might be translated into English as “maybe would unless”” (ibid. 440).

(5.26) Matses (Panoan; Fleck 2003, 440)

- (a) *cho-nunda*
come-POT.3
'(He) might come.'
- (b) *nisi-n pe-nushe tapucute ta-siuid-∅*
snake-ERG bite-POT.3 shoe foot-put.on.clothes-IMP
'Put on boots because (if you don't) a snake might bite you.'

Yurakaré is a language where the POTENTIAL marker can also appear with past reference about some action that could have happened:

(5.27) Yurakare (unclassified; Van Gijn 2006, 199)

- ti-yam pura ka-mal-ta-O samma ku-ti=la*
1SG-sibling Pura 3SG-go.SG-POT-3 water 3SG.CO-say:1SG.SBJ=VAL
'I thought the water may have taken my sister Pura.'

In Kwaza, the POTENTIAL suffix *-tsy* usually has future reference, but this is a less certain future than the FUTURE suffix *-nã* refers to. It can co-occur with the FUTURE marker and then

signals a remote FUTURE (the FUTURE marker itself does not specify any remoteness) (Van der Voort 2004, 607).

(5.28) Kwaza (unclassified; Van der Voort 2004, 607)

- (a) *tālo-**nā**-tse* / *tālo-da'mỹ-tse*
 angry-FUT-DECL / angry-want-DECL
 '(The dog) is becoming angry.'
- (b) *tā'lo-**tsy**-tse*
 angry-POT-DECL
 '(Don't tease the dog or) it will get angry.'
- (c) *ze'z̥ju-dy-rj̥* *oja-**nā**-da-**tsy**-tse*
 Zezinho-POSS-CD:area go-FUT-1-POT-DECL
 'I will go to Zezinho (one of these days).'

Unlike in Kwaza, where the co-occurrence of the POTENTIAL and FUTURE suffixes expresses a remote FUTURE, the POTENTIAL suffix (on its own) in Desano expresses an immediate FUTURE (Miller 1999, 82).

The Baure POTENTIAL particle derives from the verbal root *toeri* 'do how', as in 'How do you do the fishing?' (Danielsen 2007, 294-295). Unlike in other languages, the Baure POTENTIAL marker does not predominantly have FUTURE reference. It seems that the Baure INTENTIONAL markers are preferred to express FUTURE (Baure does not have a marked FUTURE), and that the POTENTIAL *toeri* can occur with both FUTURE and PAST reference. It was not found in conditional/ counterfactual clauses.

(5.29) Baure (Arawakan; Danielsen 2007, 295)

- (a) *rekič̥in* **toeri** *rišim?*
 when POT 3SG.F=arrive
 'When may she arrive?'
- (b) *ni=pihik* **toeri** *ne'*
 1SG=pass POT here
 'I might pass here.'

The Yaminahua POTENTIAL suffix *-tiro* cannot co-occur with Tense marking at all (Faust & Loos 2002, 125).

When a POTENTIAL marker occurs in PAST contexts it can express counterfactual clauses, e.g. in Awa Pit, which has two suffixes (positive and negative POTENTIAL). They can occur in the apodosis of counterfactual clauses (but cannot co-occur with Tense marking) (Curnow 1997, 242f.). In Desano, the POTENTIAL suffix *-bu* in combination with the recent PAST suffix *-a* marks a counterfactual. In Mekens, the POTENTIAL particle *eteet* co-occurs with the IRREALIS particle *pegat* in a counterfactual clause.

- (5.30) Desano (Tucanoan; Miller 1999, 159)

deko bēřē-biri-kī ari-bu-a-yū-bā
 water fall-NEG-SR come-CNTRFC-ASM-3PL
 ‘If it hadn’t rained, they would have come.’

- (5.31) Awa Pit (Barbacoan; Curnow 1997, 243)

akki pana-t=na, izh-sina=ma
 here be:standing(IPFV.PTCP)-CNTRFC=TOP see-POT=TEMP
 ‘If he were here, we could meet.’

- (5.32) Mekens (Tupían; Galucio 2001, 71)

kiri=eri=ēp ka-t te te se-poetop
 child=ABL=really go/come-PST truly FOC 3c-knowledge
eat pegat eteet
 acquire IRR POT
 ‘If it had been really since childhood, then I would have learned.’

In Kwaza, Awa Pit, and Warao the POTENTIAL markers can also refer to a general ability to carry out the action in question, either because circumstances allow it or because of the speaker’s personal ability. When the Kwaza POTENTIAL suffix *-tsy* occurs in combination with negation, it refers to inability (Van der Voort 2004, 610).

Awa Pit has a positive and a negative POTENTIAL marker (*-sina* and *-satshi*). When they occur in a clause unmarked for person, they express a general ability, when they occur in a clause marked for person, they express more specific possibility or ability (Curnow 1997, 242f.). The Warao POTENTIAL suffix *-komo* expresses ability as well.

- (5.33) Awa Pit (Barbacoan; Curnow 1997, 242)

- (a) *maza caballo maza carro tayη-sina*
 one horse one car pull-POT
 ‘A single horse can pull a car.’
- (b) *nyampi=kasa pishkatu ki-sina-y*
 hook=with fish(1) fish(2)-POT-N.LPM
 ‘You (the addressee) can fish with a hook (because the river is up).’

- (5.34) Warao (unclassified; Romero-Figueroa 1997, 102)

he nahobu-komo nebu koho-ya naru-a-e
 crab find-POT young.man river.mouth-ALL go-PNC-PST
 ‘The young men went to the river mouth. They can find crabs (there).’

To summarize, POTENTIAL markers in the sample can occur in the apodosis in counterfactual clauses (e.g. in Desano, Awa Pit, Mekens), signal immediate (Desano) or remote (Kwaza)

FUTURE, or ability (Warao, Awa Pit). POTENTIAL markers can be used in Matses and Nasa Yuwe for POLITE COMMANDS and in Kwaza for permission; they do not co-occur with COMMAND marking as specified in section 5.3.

The expectations from 5.5.1 are fully met. POTENTIAL markers are found with both FUTURE and PAST reference, in interrogative, negative, conditional, and IRREALIS clauses. POTENTIAL markers in a NONPAST context refer to actions that are in the realm of possibility and therefore by default questionable; although they have not been found in conjunction with Epistemic markers it is clear that they include inherent uncertainty.

Bybee et al. (1994, 177-178) propose a grammaticalization path from ability (from mental and physical to general) over root possibility to permission. Ability refers to “the existence of internal enabling conditions” and root possibility to “general enabling conditions”. The POTENTIAL marker in Awa Pit refers to ability (5.33a) (that the horse is able to pull the car is conditioned by its internal values such as strengths), but also to root possibility (5.33b) (the possibility to fish with a hook is conditioned by speaker-external factors such as that the river is up). The POTENTIAL marker in Warao likewise points towards root possibility (finding crabs is conditioned by speaker-external factors, such as that crabs generally can be found at the river).

Sadly, the sources of the POTENTIAL markers in the sample remain largely unspecified. The POTENTIAL suffix in Movima (-*nira*) is possibly composed of the verbalizing suffix -*ni* and the semantically empty element -*ra* (Haude 2006, 432).

5.6 PURPOSIVE

5.6.1 Definition

Although this study is carried out from a semantic perspective, in order to define PURPOSIVE it is necessary to comment on the syntax of purpose clauses. A speaker can indicate that the action of the matrix clause is done for a purpose, like “X does Y in order to achieve Z”: “[p]urpose clauses are part of complex sentences which encode that one verbal situation, that of the matrix clause, is performed with the intention of bringing about another situation, that of the purpose clause” (Schmidtke-Bode 2009, 20). The PURPOSIVE marker in these cases is the one carrying the meaning of “in order to”, “with the intention” to, etc. The term “intention” will be avoided as much as possible in this section in order to avoid confusion with the INTENTIONAL. This also goes to show that PURPOSIVE and INTENTIONAL overlap, i.e. behind each purpose is the intention to do X. Both INTENTIONAL and PURPOSIVE markers refer to an event situated in the future, both are goal-oriented (i.e. directed toward a prospective target state or event that is to be deliberately achieved): “Purpose relations link two [clauses] one of which (the main one) is performed with the goal of obtaining the realization of another one (the dependent one)” (Cristofaro 2003, 157). But while INTENTIONAL is usually expressed in a simple main clause, PURPOSIVE is usually expressed in a complex clause. They constitute the same semantics but differ in construction complexity. The purpose clause is usually subordinate, which means that a PURPOSIVE can be a subordinator (or a conjunction). Not every subordinator is a PURPOSIVE, however, only when its primary

meaning is purpose is it coded as *PURPOSIVE*, and not every *PURPOSIVE* necessarily occurs in subordinated clauses (as will be seen below). A prototypical *PURPOSIVE* marker is found in Cavineña (=ishu):

(5.35) Cavineña (Tacanan; Guillaume 2008, 707)

camion nubi=**ishu**=tuna-ja=tu e-dijo bajeje-ti-chine
 truck enter=PRP=3PL-DAT=3SG NPF-path prepare-TMPR-REC.PST
 ‘They went there to arrange the path so that the trucks can enter.’

An example for a marker that expresses *PURPOSIVE*, but is not a *PURPOSIVE* marker in the sense of the present definition is the suffix *-m* in Mapuche. This instrumental suffix obligatorily occurs with other markers and only expresses *PURPOSIVE* in combination with the *IRREALIS* marker *-a* (Smeets 2008, 206). The fact that *PURPOSIVE* is not its primary meaning and that a *PURPOSIVE* meaning is only given in combination with another marker, disqualifies *-m* as *PURPOSIVE* in my sample.

(5.36) Mapuche (Araucanian; Smeets 2008, 209)

fey amu-y ñi ngilla-ka-ya-**m**
 she go-IND-3 POSS3 buy-FCT-IRR-IVN
 ‘She went in order to go shopping.’

5.6.2 Distribution in the sample

PURPOSIVE is marked by 24 out of the 63 languages in the sample:

Tariana, Tsafiki, Trumai, Awa Pit, Hixkaryana, Tiriyó, Ika, Embera, Mosetén, Mamaindé, Imbabura Quechua, Emérillon, Yaminahua, Cavineña, Karo, Nheengatú, Cocama-Cocamilla, Leko, Kwaza, Munich, Warao, Shipibo-Konibo, Miraña, and Matses.

Formal marking includes suffixation, enclitics, and particles, with mostly suffixes.

PURPOSIVE markers occur in indicative, but also in *COMMAND* clauses and are in some languages organized in paradigms according to whether the action of the clause is referred to by a motion verb, and whether the purpose clause has the same or a different subject than the main clause. The following paragraphs outline these functions.

The language with the most elaborate *PURPOSIVE* system is Mamaindé. It employs distinct suffixes for positive and negative purpose clauses and further divides positive purpose clauses into same subject and different subject clauses, each with different suffixes. Imbabura Quechua also has two distinct suffixes for same subject and different subject (*-chun* and *-ngapaj*, respectively) (Cole 1985, 63). The suffix *-ngapaj* consists of the *FUTURE* (3rd person) suffix *-nga* and *-paj* ‘for’ (ibid. 37). In comparison, Huallaga and Cuzco Quechua do not mark *PURPOSIVE*. Cole (1985, 37f.) mentions that the specific pattern of Imbabura Quechua

subjunctive clauses (of purpose) is uncommon in the Quechua dialects. Purpose in Huallaga Quechua is instead expressed by special construction including the suffix *-paq*, which has various functions (Weber 1989, 204), a possessive suffix, and the suffix *-na*, which is a substantivizing subordinator (ibid. 205).

(5.37) Mamaindê (Nambikwaran; Eberhard 2009, 535, 536)

- (a) *jalik-tu waun-k^hanîn-tu tu-te?nta? joha-ten-a?-wa*
 necklace.FNS red-NCL.round-FNS get-CN.PRP.SS trade-DES-SBJ1-PRS-DECL
 ‘I want to trade in order to get red necklace beads.’
- (b) *hâi nusa-hai?ka nakat-si?*
 he POSS1.PL-word hear/understand-CN.PRP.DSBJ
nawih-ta? set-a?-Ø-wa
 teach-CN.and.SS speak-SBJ1-PRS-DECL
 ‘For him to understand our language, I am teaching and speaking.’

There are two other constructions for purpose motion clauses and negative *PURPOSIVE*. In Leko, the suffix *-to* marks *PURPOSIVE* clauses with motion verbs. Another strategy to express *PURPOSIVE* is the benefactive marker *-moki*,¹⁴ but it is unclear whether these two markers constitute a motion/ nonmotion *PURPOSIVE* system (Van de Kerke p.c.).

(5.38) Leko (unclassified; Van de Kerke 2009, 324)

- yo-moki choswai wesra yu-nun-a-ka-te aycha hal-to*
 1-GEN woman Guanay 1.OBJ-command-PFV-AUX-DECL meat buy-PRP
 ‘My wife sent me to Guanay in order to buy meat.’ (orig. ‘Mi esposa me ha mandado a Guanay a comprar carne.’)

Yaminahua also marks same subject/ different subject *PURPOSIVE* clauses with different markers. The suffix *-nō* is used when the subjects of main and subordinate clauses differ. The suffix *-xiki* is used for same-subject clauses and probably originates in the combination of the suffix *-xi* ‘para’ and *-ki* ‘future’ (Faust & Loos 2002, 155).

Interestingly, Tsafiki, a Barbaocoan language that is not related to Quechuan but geographically nearby, also has a *PURPOSIVE* marker *-chun* (Dickinson 2002, 132).¹⁵ This does not necessarily point toward a relationship between Tsafiki and Imbabura Quechua, however. In Tsafiki, the combination of the *IRREALIS* suffix *-chu* and the stative suffix *-n* occurs quite frequently (cf. Dickinson 2002). Formally, this may be the origin of a reinterpreted *PURPOSIVE* suffix *-chun*, although the exact semantic development is not clear.

¹⁴ Benefactive markers expressing *PURPOSIVE* are found all over the world (cf. Schmidtke-Bode 2009, 90).

¹⁵ Matses has an adverbial clause subordinator *-shun* ‘after’, ‘when’ (Fleck 2003, 1117).

(5.39) Tsafiki (Barbacoan; Dickinson 2002, 132)

junni aman in=la=ka man=fi-chu-n aman wa paila
 then now 3p1=PL-ACC again=eat-IRR-STA now big pan
da-kari-man-ti-e.

cook-cause:generic-SIT-REPO-DECL

‘Then they say she heated up a big pan to eat them again.’

As was seen above, Leko has a marker specifically for PURPOSIVE with motion verbs, whereas non-motion verbs purpose clauses are marked by the benefactive suffix *-moki*. Other languages also employ different markers according to motion or non-motion verbs. For example, Cavineña has a general PURPOSIVE enclitic *=ishu* that can occur with all verb types (Guillaume 2008, 707), but an enclitic *=ra* specifically for PURPOSIVE with a controlled motion in the matrix clause (ibid. 715). In Tiriyo, the MOTION PURPOSIVE suffix *-se* is used when the matrix verb is a motion verb (usually ‘go’ or ‘come’) (Meira 1999, 569), with non-motion verbs the PURPOSIVE is expressed by a different construction (ibid. 331). Hixkaryana has a suffix similar to the Tiriyo MOTION PURPOSIVE: *-se/ -xe* that also marks MOTION PURPOSIVE (Derbyshire 1985, 40f.), But Hixkaryana also has a PURPOSIVE particle that is apparently not restricted to motion verbs: *hori* (ibid. 39).

(5.40) Cavineña (Tacanan; Guillaume 2008, 715, 717)

(a) *tudya e-kwe e-mama e-kwe e-tata*
 then 1SG-GEN 1-mother 1SG-GEN 1-father
jadya kwa-kware nawi=ra

and go-REM.PST bathe=MOT.PRP

‘My father and my mother went to bathe.’

(b) *i-ke babi=ra kwa-ya*
 1SG-FM hunt=MOT.PRP go-IPFV

‘(I told my wife:) I’m going hunting!’

(5.41) Tiriyo (Cariban; Meira 1999, 331)

epë-e wi-të-e
 bathe-PRP 1S_A-GO:PRS.IPFV-CERT

‘I am going (to the river) to bathe.’

Schmidtke-Bode (2009, 95), in his study of purpose clauses uses a global sample, discovered that MOTION PURPOSIVE markers are common and “not confined to particular geographical [macro]areas or language families”. This can be confirmed for the SAILS in the sample. In the present sample, five languages explicitly code MOTION PURPOSIVE (Leko, Hixkaryana, Tiriyo, Cavineña, and Embera). The present data also show that PURPOSIVE markers occur mostly in same-subject relations between main and subordinate clause, except for the cases where there is a special different-subject PURPOSIVE marker. The close connection of

PURPOSIVE to motion and same-subject involvement is natural, as Cristofaro (2009, 157) points out: “In fact, most of the data on purpose relations concern the purpose of motion [...]. In motion purpose relations, an entity goes somewhere in order to obtain the realization of a certain SoA [SoA=state of affairs], and it is usually assumed (though not logically entailed) that this realization is brought about by the entity itself”.

Some of the languages exhibit PURPOSIVE markers in command clauses. These do not have a matrix clause-subordinate clause structure, but simply a single clause with the verb marked with a COMMAND marker. This is evidence for the fact that the semantics of PURPOSIVE transcend a certain syntactic construction. The semantic connection between PURPOSIVE and COMMAND is not hard to explain. As Schmidtke-Bode (2009, 145) states, a “phenomenon pertaining to illocutionary force is that purposive constructions lend themselves well to being formulated as directives”. In Tiriyo, both HORTATIVE and PURPOSIVE markers can co-occur (Tiriyo PURPOSIVE marker *-se* (for PURPOSIVE in motion clauses). In Mosestén, the PURPOSIVE particle *paj* itself can express an indirect COMMAND (Sakel 2004, 439)).¹⁶

(5.42) Tiriyo (Cariban; Meira 1999, 331)

npa epë-e
 HORT bathe-PRP
 ‘Let’s go bathe!’

(5.43) Mosestén (Mosenan; Sakel 2004, 439)

jam-ra’ mi’in bojw-i paj mama’ bojw-il
 NEG-IRR 2PL go.up-VM-M.SBJ PRP father go.up-VM-M.SBJ
 ‘Don’t you climb up, let father climb up!’

The Kwaza PURPOSIVE marker *-te* is most common with first person singular (Van der Voort 2004, 412) and is possibly a nominalizer. In contrast, the Cocama-Cocamilla PURPOSIVE marker *=ra* is homophonous with the verbalizer *-ra* (Vallejos Yopán 2010, 295, 385).

Tariana is the only language in the sample that encodes both PURPOSIVE and Tense/Evidentiality in cumulative morphemes. The VISUAL or IMMEDIATE PURPOSIVE is marked by *-karu*, and the NONVISUAL or DISTANT PURPOSIVE by *-hyu* (Aikhenvald 2003, 393).

(5.44) Tariana (Arawakan; Aikhenvald 2003, 393)

(a) *wasā tarada-peni wehta-karu wa*
 let’s.go alive-PL:AN 1PL+take-VIS.PRP 1PL+go
 ‘Let’s go and see the living ones (fish) (right now).’

¹⁶ Apparently, Paumari has a grammaticalized PURPOSIVE command construction (Schmidtke-Bode 2009, 145).

- (b) *taPda-peni duhta-hyu du-a-pidana*
 alive-PL:AN 3SG.F+take- NVIS. PURP 3SG.F-go-REM.PST.REPO
 ‘She went to get the living ones (she may not see them).’

As was already pointed out in the introduction, the distinction between *PURPOSIVE* and *INTENTIONAL* lies mostly in syntactic complexity, in that a *PURPOSIVE* always expresses “purposive relations between two states of affairs in the world” (Schmidtke-Bode 2009, 30). Nevertheless, we have also seen that *PURPOSIVE* markers can occur in different constructions, as in command clauses. It is common that languages with either only *INTENTIONAL* or *PURPOSIVE* markers or both have one of these markers expressing the semantics of the other as a secondary meaning. For example, Matses has a range of five *PURPOSIVE* markers, two of which can also indicate *INTENTION* (all of the markers have the secondary temporal meaning ‘before’ (Fleck 2003, 1110 ff.)).¹⁷

To summarize, languages can have a single *PURPOSIVE* marker that only refers to motion clauses (e.g. Tiriyo, Hixkaryana, Cavineña), *PURPOSIVE* markers that do not make a distinction between motion or non-motion clauses (e.g. Hixkaryana, Cavineña), *PURPOSIVE* markers specifying for same-subject and/ or different-subject purpose clauses (Mamaindê, Imbabura Quechua, Yaminahua, Awa Pit), and *PURPOSIVE* markers expressing directives in purpose clauses (e.g. Tiriyo, Mosestén). Mamaindê is the only language specifying negative and positive *PURPOSIVE* with different markers.

5.7 FRUSTRATIVE

5.7.1 Definition

A *FRUSTRATIVE* refers to an event that did not have the expected outcome or was finished unsuccessfully. The action can be left unfinished, or be finished but not as expected, or be done in vain. It involves emotive frustration on the part of the speaker, but not necessarily so. It is not an *INCOMPLETIVE*, which just states that an action is not finished, regardless of whether the outcome was expected or desired. One could say that semantically a *FRUSTRATIVE* marker can be an *INCOMPLETIVE* with added frustration in those cases where the action is not finished, but this is only a part of *FRUSTRATIVE* meaning. Actions may very well be finished, which prohibits *INCOMPLETIVE* meaning, but not with the desired outcome. A typical example of a *FRUSTRATIVE* marker is the clitic =*tha* in Tariana:

- (5.45) Tariana (Arawakan; Aikhenvald 2003, 380)
nuha nu-sata-tha-na nhuma
 I 1SG-ask-FRUST-REM.PST.VIS 1SG+hear
 ‘I did try in vain to ask (the pilot why he did not let us go).’

¹⁷ The intentional dependent clause (same-subject) can express purpose in Aguaruna (Overall 2007, 504).

As can already be observed from the above paragraph, a FRUSTRATIVE can have several shades of meanings. All have in common that an intended result or goal is not reached, whether the action was finished or not. The resulting state can include frustration and dissatisfaction on part of the speaker.

5.7.2 Distribution in the sample

No less than 23 languages in the sample mark FRUSTRATIVE morpho-syntactically:¹⁸

Apurinã, Jarawara, Tariana, Paresi, Miraña, Hixkaryana, Tiriyó, Aguaruna, Timbira, Hup, Dâw, Mapuche, Mosetén, Cavineña, Cubeo, Desano, Rikbaktsa, Mekens, Kamaiurá, Nheengatú, Movima, Kwaza, and Puinave.

Forms range from suffixes over clitics and particles to one auxiliary (Dâw). The following section gives an overview of the FRUSTRATIVE markers and its applications, starting with contrary-to-fact environments, unusual meaning extensions, and origins.

That FRUSTRATIVE marking is common in counterfactual statements has already been noted by Overall (2007, 386) and can be confirmed for the present data set. A selection of languages with FRUSTRATIVE markers in counterfactual contexts is given below.¹⁹

When the FRUSTRATIVE marker *-ma* in Apurinã co-occurs with the FUTURE marker it has the meaning of a counterfactual clause (Facundes 2000, 534). It attaches to verbal and nonverbal bases, but always refers to the action of the clause. The same is possible with the Hixkaryana FRUSTRATIVE particle *haryhe*. In Hup, the FRUSTRATIVE marker *yã'h* often occurs in counterfactual clauses which are marked by *tã'?*, but it is also possible to have a counterfactual clause without the counterfactual but with FRUSTRATIVE marker. In Mapuche, the FRUSTRATIVE suffix *-fu* can occur in conditional clauses that are not counterfactual but hypothetical. The following examples illustrate the usage of the FRUSTRATIVE markers discussed above.

(5.46) Apurinã (Arawakan; Facundes 2000, 404, 535)

- | | | | |
|-----|---|-------------|-------------------|
| (a) | <i>nu-nuta-pe-ka-ma-ru</i> | <i>kona</i> | <i>n-apoka-ru</i> |
| | 1SG-search-PFV-PASS-FRUST ²⁰ -3M.OBJ | not | 1SG-find-3M.OBJ |
| | 'I searched it/ him but I didn't find it/ him.' | | |

¹⁸ The status of the FRUSTRATIVE suffix *-kean* in Shipibo-Konibo is unclear. Valenzuela (2003, 284) does not give any examples; the usage of *-kean* in Valenzuela (1997) points more toward INCOMPLETEIVE.

¹⁹ The FRUSTRATIVE could not be found in counterfactual clauses in Mosetén (in Sakel 2004). In Cavineña, no counterfactual clauses were found (in Guillaume 2008); the FRUSTRATIVE = *datse* does not occur in conditional clauses. In Cubeo, the FRUSTRATIVE suffix *-bu* has not been found in conditional (counterfactual or not) clauses (in Morse & Maxwell 1999).

²⁰ In the original text the gloss FRUST is missing; instead *-ma* is erroneously glossed as 'PASS'.

- (b) *nhi-nhipoko-ta-ma-ko kona awa-ru nhipoko-ru*
 1SG-food-VBLZ-FRUST-FUT not be-3M.OBJ food-UNPOS
 ‘I would eat but/ if there’s/ were no food.’

(5.47) Hixkaryana (Cariban; Derbyshire 1985, 253)

- (a) *nekaimyatxkon haryhe ti*
 they-were-climbing FRUST HEA
 ‘They were trying to climb (but didn’t succeed).’
- (b) *ihokhura ehtoko, enamri tho haryhe*
 without-child if-his-being burying-of-him DEVL FRUST
 ‘If he had had no children, he would have been buried.’ (As it was he was cremated.)

(5.48) Hup (Nadahup; Epps 2008, 614, 623)

- (a) *hũtǎh ʔǎh ní-tǎn, ʔǎh way-d’oʔ-tǎʔ-ǎy yǎh*
 bird 1SG be-COND 1SG leave-take-CNTRFC-DYN FRUST
 ‘If I were a bird, I would fly.’
- (b) *tih d’óʔ-óy yǎh=mah yúp, tǎʔǎy-ǎw-ǎp*
 3SG take-DYN FRUST=REPO that.ITG woman-FLR-DEPM
 ‘She would have taken him (as her husband), the woman.’

(5.49) Mapuche (Araucanian; Smeets 2008, 232)

- amu-a-l-m-i wekun, makuñ-tu-nge*
 go-IRR-COND-2-SG outside, coat-VBLZ-IMP2SG
chafo-a-fu-y-m-i wütre-mew
 catch.a.cold-IRR-FRUST-IND-2-SG cold-INST
 ‘If you go outside, put on a coat. You might catch a cold in view of the cold.’

The FRUSTRATIVE marker *-fu* in Mapuche shows meanings not usually related to FRUSTATIVE. It can denote that the action is over, ended, which is the case for all frustrative markers, but also that the evaluated outcome is not negative, which is not prototypical (cf. examples (6) and (10) in Smeets 2008, 231-232). It has previously been analyzed as a PAST Tense marker (likewise, the IRREALIS suffix *-a* has been analyzed as FUTURE Tense marker (ibid. 234)), but Smeets (ibid.) demonstrates that the PAST meaning derives from the Modal values of *-fu* rather than being the primary meaning. The FRUSTRATIVE in Mapuche is also closely connected to the IRREALIS marker *-a*, with which it often co-occurs (cf. example 5.49), so in previous publications analyzed the sequence *-a-fu-* as one morpheme *-afu*.

The FRUSTRATIVE particle in Movima could not be found in counterfactual clauses (in Haude 2006), but interestingly counterfactual clauses are marked by the particle *di* which also appears in the FRUSTRATIVE particle *didi* (several Modals have this form: FRUSTRATIVE *didi*, hypothetical *di*, optative *dis*, counterfactual *disoy* (Haude 2006, 530)).

In Tiriyo, the FRUSTRATIVE marker *re(pe)* refers to an action that was not planned but almost happened anyway, to the displeasure of the speaker. It can co-occur with the DESIDERATIVE on the same verb. The Cubeo FRUSTRATIVE suffix *-du* also co-occurs with the DESIDERATIVE, as does the Mekens FRUSTRATIVE particle *etaop* (cf. Galucio 2001, 151, 5b).

(5.50) Tiriyo (Cariban; Meira 1999, 479)

j-anota_re
1S_o-fall:PRS.PFV_FRUST
'I almost fell.'

(5.51) Cubeo (Tucanoan; Morse and Maxwell 1999, 29)

dō-I ʔ-RE õpõ-boa-ji-Ri-du-RExa-kaki
that-LOC 3M.SG-OBJ explosion-kill-DES-GER-FRUST-N.REC.PST/PRS.HAB.1M.SG
'Over there, I wanted to shoot him (the tapir) (but didn't).'

In Hup, the action marked with FRUSTRATIVE *yã'h* can even be successfully completed, but something that happened after that had frustrating effects on the previous action. In (5.52), a necklace was created, although it later got lost or stolen.

(5.52) Hup (Nadahup; Epps 2008, 620)

tit w'ã-õt ʔãh cuh-ʔe?-yãh-ãh
string long-OBQ 1SG string-PFV-FRUST-DECL
'I strung (the beads) on a long string (in vain).'

The FRUSTRATIVE particle *jepe* in Kamaiurá has two functions that rarely occur with FRUSTRATIVE meaning: it can indicate that the marked verbal action is about to happen when it co-occurs with the particle *rak* 'atestado'; and it can indicate a fictitious event (Seki 2000, 96). The FRUSTRATIVE suffix *-hama* in Jarawara also can refer to a pretended action (Dixon 2004, 171), as well as the Kwaza suffix *-le* (ambiguous between FRUSTRATIVE and INCOMPLETEIVE; see below).

(5.53) Kamaiurá (Tupían, Guaraní; Seki 2000, 96)

- (a) *o-mano jepe rak*
3-morer FRUST ATTESTED
'He almost died.' (orig. 'Ele quase morreu.')
- (b) *o-jae'o jepe*
3-cry FRUST
'He pretends to cry.' (orig. 'Ele finge que chora.')

(5.54) Jarawara (Arawan; Dixon 2004, 171)

- o-ko-make-hama o-na-hara o-ke
 1SG.SBJ-in.motion-FOLLOWING-FRUST 1SG-AUX-IMM.PST.EYE.F 1SG-DECL.F
 ‘I pretend to run (say: someone else was running, asked me to run with him, I said I would, started to run, but then stopped without saying I was stopping).’

The FRUSTRATIVE suffix *-le* in Kwaza is a good example for a marker that is ambiguous between FRUSTRATIVE and INCOMPLETEIVE. It marks actions that are not completed (and even interrupted), or did not start at all (Van der Voort 2004, 432), but can also refer to an action that was intended but turned out to be impossible (ibid. 644).

(5.55) Kwaza (unclassified; Van der Voort 2004, 432, 641)

- (a) hy’ja-xa-**le**-ki
 fall-2-INCP-DECL
 ‘You nearly fell.’
- (b) ywy’nwỹ ‘dy=asa-’nã-da-**le**-hỹ ‘tã-i-ki
 tree cut=leave-FUT-1SP-FRUST-NMLZ hard-DECL
 ‘I was going to cut away the tree trunk, but it was (too) hard.’

How FRUSTRATIVE markers developed is certainly one of their interesting aspects. Two sources in the sample present hypotheses for Aguaruna and Hup.²¹ Overall (2007, 387) suggests that the Aguaruna FRUSTRATIVE marker *-takama* is a recently grammaticalized form and that it developed from “a periphrastic intentional construction to a subordinating intentional suffix that was always accompanied by a controlling verb indicating an unsuccessful outcome. A subsequent shift of the ‘frustrative’ meaning from the controlling verb to the subordinate verb gives us the current system in Aguaruna”.

Epps (2008, 626) suggests a grammaticalization path for the FRUSTRATIVE marker *yã’h* of the verb with the same form meaning ‘order, compel, request’, which is more straightforward than what Overall presents for Aguaruna. These two developments point toward language-internal change, rather than contact, but the FRUSTRATIVE as a functional category can apparently be found as a borrowing in Mawayana: Carlin (2007, 313) argues that the functional category FRUSTRATIVE is borrowed into Mawayana (Arawakan) from Cariban, where it is obligatory. Both Cariban languages in the sample have FRUSTRATIVE markers (Hixkaryana *haryhe*, Tiriyó *re(pe)*), though they do not seem to be related to Mawayana *-muku* (etymology unknown). Carlin bases her argument for the borrowing of the functional category on the fact that it “has, for the most part, exactly the same morphosyntactic properties as the Cariban frustrative, that is, it can be marked on the major verb classes, and it carries the same meaning” (ibid. 327). If this is indeed the case, a

²¹ Awetí, a Tupí language not included in the sample, may have developed the IRREALIS marker *tutepe* from the FUTURE marker *tut* and the FRUSTRATIVE marker *tepe* (Drude 2008, 92).

worthwhile future investigation of similar contact scenarios for FRUSTRATIVE marking is highly desirable.

For a discussion about FRUSTRATIVE as an areal characteristic see section 5.10.

5.8 DESIDERATIVE

5.8.1 Definition

A DESIDERATIVE marker pertains to a wish or desire of the subject of the clause to carry out an action, to acquire an object or skill, or for a certain event to take place. This can be the first person (speaker), but also any other kind of subject, as in Apurinã:

(5.56) Apurinã (Arawakan; Facundes 2000, 316)

- (a) *nhi-nhika-ene-ta-ru*
1SG-eat-DES-VBLZ-3M
'I wanted to eat it.'
- (b) *hākiti akatsa-ene-ta-ru*
jaguar bite-DES-VBLZ-3M
'The jaguar wanted to bite him.'

Other terms than DESIDERATIVE found in the literature, which sometimes denote the same meaning and sometimes have a slightly different one, are optative and volitive. Sometimes, so-called exhortatives are used to express wishes, but since those forms are usually restricted to first person COMMANDS, they are coded as HORTATIVE (see section 5.9).

5.8.2 Distribution in the sample

In the present sample of 63 languages 39 morpho-syntactically mark DESIDERATIVE:

Apurinã, Jarawara, Aymara, Awa Pit, Tsafiki, Hixkaryana, Panare, Tiriyo, Aguaruna, Mocoví, Pilagá, Bororo, Dâw, Hup, Mamaindê, Sabanê, Matses, Shipibo-Konibo, Yaminahua, Huallaga Quechua, Imbabura Quechua, Cavineña, Cubeo, Desano, Sateré-Mawé, Karitiana, Mekens, Emérillon, Kamaiurá, Tapiete, Leko, Movima, Yurakaré, Kanoê, Kwaza, Trumai, Munich, Puinave, and Warao.

The form of marking ranges from suffixation over particles to auxiliaries with a clear preference for suffixes.

DESIDERATIVE markers are frequently found to denote meanings that are related to desire in both Modality and Tense (besides their prototypical meaning). The following section explores the DESIDERATIVE having INTENTIONAL, FRUSTRATIVE, PURPOSIVE, FUTURE, and directive functions, as well as its distribution patterns and semantic restrictions.

DESIDERATIVE and Modality

Not surprisingly, DESIDERATIVE markers can extend to Modal meanings that are related to the sense of desire, as e.g. INTENTIONAL. Two languages have markers that are ambiguous between DESIDERATIVE and INTENTIONAL: Jarawara and Mamaindê. The DESIDERATIVE/INTENTIONAL suffix *-bone* in Jarawara can be used for both meanings, as can the suffix *-ten* in Mamaindê (which can in addition refer to immediate FUTURE). The DESIDERATIVE suffix *-chike* in Tsafiki only occurs in subordinated purpose clauses. This suffix and the PURPOSIVE suffix *-chun* both encode purpose, but *-chike* adds a sense of desire (Dickinson 2002, 132).

The DESIDERATIVE suffix in Hixkaryana can imply effort on the part of the speaker (Derbyshire 1985, 239). This differs from traditional DESIDERATIVES, where the desired action does/ did not (yet) take place, but the action is/ was carried out unsuccessfully with a sense of 'to try'. This relates to FRUSTRATIVE meaning. Hixkaryana also employs the DESIDERATIVE marker *ʃe* (Derbyshire 1999, 58).

- (5.57) Hixkaryana (Cariban; Derbyshire 1985, 239)

ti-niki-txahke

ADVZ-sleep-DES

'wanting to sleep', 'trying to sleep'

- (5.58) Hixkaryana (Cariban; Derbyshire 1999, 59)

o-ʃe w-ehf-aha

2-DES 1SBJ-be-NPST

'I love you.'

Desiderative and speech-acts

Matses has three ways to express desire: with the verb *bun* 'want' with a complement clause as object (Fleck 2003, 1034), with the abilitative (willingness) suffix (ibid. 1053), or with the suffix *-pashun* (ibid. 434). The first two only apply to same-subject situations. When *-pashun* is used, the speaker expresses his desire that somebody else do something, i.e. this suffix denotes a directive. It is not a command, nor even a weak command, as there is virtually no force behind the proposition, but it is directed towards an addressee (that may even be absent). Aymara has a paradigm of person and DESIDERATIVE suffixes. It is used for desire, admonitions, and recommendations (Hardman 2001, 114). In Kwaza, when the subject/agent of the clause with the DESIDERATIVE suffix *heta* is not a first person, it also has directive function.

- (5.59) Matses (Panoan; Fleck 2003, 434)

(a) *acte bacush ac-pashun*

water foam drink-DES

'I wish he would drink beer.'

- (b) *acte bacush mibi ac-pashun*
 water foam 2ERG drink-DES
 'I wish you would drink beer.'

- (5.60) Kwaza (unclassified; Van der Voort 2004, 416)
pērējā-he'ta-da-(le)-ki
 speak-DES-1SG-FRUST-DECL
 'I would like him to talk.'

Desiderative and Future

The Aguaruna DESIDERATIVE suffix is homophonous with the definite FUTURE suffix (-*tata*); the latter is more grammaticalized. The suffix may be derived from the immediate FUTURE suffix -*ta* plus the verb *tu* 'say' (Overall 2007, 355f.). The Hup auxiliary *tu* also marks immediate FUTURE, and the Mamaindê suffix -*ten*, which is ambiguous between DESIDERATIVE and INTENTIONAL, can also have immediate FUTURE meaning. It cannot CO-OCCUR with the FUTURE markers and is indeed an alternative FUTURE marker in its own right (Eberhard 2009, 420).

- (5.61) Mamaindê (Nambikwaran; Eberhard 2009, 419, 429)
- (a) *?iun-ten-a?-Ø-wa*
 sleep-DES-1SBJ-PRS-DECL
 'I intend to sleep.'
- (b) *wa-māin-tu tɔ-ten-lat^ha-Ø-wa*
 POSS2-pet-FNS die-DES-3SBJ-PRS-DECL
 'Your pet will die.'

The DESIDERATIVE suffix -*sicha'kwa* in Movima can also refer to actions that are about to occur when the (only) participant of the clause is an undergoer (Haude 2006, 445). One instance of the DESIDERATIVE suffix -*nta* in Yurakaré also exhibits an 'about to' meaning:

- (5.62) Yurakaré (unclassified; Van Gijn 2006, 200)
mashi-nta-Ø
 rain-DES-3
 'It wants to rain (i.e.: it looks like it's going to rain).'

One DESIDERATIVE marker was found which remotely relates to the Aspectual INCOMPLETE: The Karitiana DESIDERATIVE suffix -*wak* marks events that nearly happened (Everett 2006, 280), therefore adding a sense of incompleteness as well as FRUSTRATIVE.

(5.63) Karitiana (Tupían; Everett 2006, 280)

- (a) *ĩn i-pítʔi-wak*
 1SG INTR-eat-DES
 ‘I want to eat.’
- (b) *i-taka-tat-awak ĩn*
 1SG.ABS-SAP-go-DES 1SG
 ‘I almost went.’

Bodily functions/urges, and weather predicates

The Kanoê DESIDERATIVE suffix usually occurs with bodily functions, but not exclusively.²² Huallaga Quechua has a DESIDERATIVE suffix *-na*: which occurs on nouns (Weber 1989, 33), but also on verbs; in the latter case it occurs with bodily functions and weather predicates²³ (ibid. 170f.). The DESIDERATIVE on verbs can trigger a slight change of meaning, as seen in (5.64b) and (5.64c). Imbabura Quechua has a DESIDERATIVE suffix which is not so restricted semantically and also occurs on both nouns and verbs (cf. Cole 1985, 180f.). The DESIDERATIVE suffix *-wej* in Kamaiurá is restricted to physical desire or appetite. Any other desire has to be expressed by the DESIDERATIVE suffix *-potat* (Seki 2000, 132).

(5.64) Huallaga Quechua (Quechuan; Weber 1989, 33, 170, 171)

- (a) *yaku-na-*
 water-DES
 ‘to be thirsty, to want water’
- (b) *chiwlla-na-*
 to.urate-DES
 ‘to need to urinate’
- (c) *qeshya-na-sha qoya-yka-n chakay tamyá-rku-r*
 sick-DES-PTCP PASS:day-IPFV-3 night rain-ASPC-ADV
 ‘All day it is clouded over, having rained last night.’

Bororo has examples with the DESIDERATIVE marker and the verb ‘urinate’:

(5.65) Bororo (Macro-Gêan; Nonato 2008, 114)

- (*i aidü re*) *i kigurudu wo*
 (1SG want ASSR) 1SG urinate DES
 ‘I want to urinate.’ (orig. ‘eu quero urinar.’)

In Movima, the DESIDERATIVE suffix *-sicha'kwa* also occurs in clauses which denote a ‘physical urge’:

²² For instance, sleep, urinate, eat, suck, speak, knock down, play, carry, lay down (cf. Bacelar 2004).

²³ The Yurakaré DESIDERATIVE suffix *-nta* also occurs with weather (rain) (cf. (5.62)).

(5.66) Movima (unclassified; Haude 2006, 446)

<i>jayna</i>	<i>n-os</i>	<i>zara:-wa,</i>	<i>in</i>	<i>choj-sicha'kwa,</i>
DISC	OBQ-ART.N.PST	wake_up-NMLZ	1INTR	urinate.DES
<i>ba</i>	<i>kwayi:-maj</i>	<i>n-os</i>	<i>po'mo:-wa</i>	
but	laziness-VLC	OBQ-ART.N.PST	get_up-NMLZ	

'As I woke up, I had to urinate, but I was too lazy to get up.'

Most of the DESIDERATIVE markers found in some way originated in forms (mostly verbs) with similar meaning. For example, the Cubeo DESIDERATIVE suffix *-jji/ -ji* derives from the stative verb *i* 'want, desire, love' (Morse and Maxwell 1999, 28). The DESIDERATIVE suffix *-potat* in Kamairú also exists as independent verb *potat* 'want, enjoy' (Seki 2000, 132). Rose (2003, 426) claims that the DESIDERATIVE suffix *-tanē ~ -tane* derives from the verb *potal* 'want' followed by the particle *ne ~ nē* 'contrastive'. The verb *potal* may also be the origin of the FUTURE suffix *-tal*. The Yurakaré suffix *-nta* may derive from INTENTIONAL *-ni* and POTENTIAL *-ta* (Van Gijn 2006, 200).

Although the DESIDERATIVE forms in the sample mostly differ very much from one another, a few languages exhibit similar forms. The strongest case of the assumption that we are dealing with the same form cross-linguistically can be made for the Quechuan, Nambikwaran and Panoan languages in the sample, but also Tupían (Guaraní). For a more detailed study of DESIDERATIVE as a characteristic of SA due to a combination of areal and genealogical effects see Mueller (2013).

5.9 COMMAND TYPES

5.9.1 Definitions

COMMANDS fall under Speaker-oriented Modality ('event modality' in Palmer 2001) and "allow the speaker to impose [...] conditions on the addressee" (Bybee et al. 1994, 179). A COMMAND is a speech act that requires at least two participants: a speaker uttering the command and an addressee who is being commanded. There is no limit to how many persons are at either the commanding or receiving end. A speaker can command by himself or as an authority on behalf of a group, and the addressee can be a specific person, a group, and even include the speaker himself. This section deals with a selection of COMMANDS that will be outlined below. At the onset, a note about terminology is necessary, however.

The terms 'COMMAND' and 'IMPERATIVE' are often used synonymously, but as Aikhenvald (2010, 1) points out, a command can be given by other means than IMPERATIVE and an IMPERATIVE does not necessarily convey a command. Here, 'IMPERATIVE' is a term restricted to second person COMMANDS, and 'COMMAND' refers to a speech-act with a force behind it that prompts an addressee to carry out the desired action. If this action is marked it receives the respective term (as discussed below). The term COMMAND is used as an umbrella-term for all of these terms.

Searle (1975) defines COMMANDS as directive speech-acts that have illocutionary force, i.e. cause the addressee to do something. This also includes advice, requests, and warnings,

besides others, which are generated by how much force there is behind the command. The traditional and prototypical COMMAND is given to a second person addressee and has force behind it as in ‘You have to finish your thesis!’, but other COMMAND types, such as POLITE COMMANDS, have lesser force. It is a topic of great discussion of where a true COMMAND ends and other directives, such as requests, wishes, or suggestions, begin, and this is even further complicated by specific social contexts. The present study acknowledges these problems, but circumvents them in trusting in the original analyses of the sources.

In the literature, COMMANDS have been given a range of terms such as ‘hortative, hortatory, cohortative, exhortatory’, and ‘jussive’. The latter is traditionally used with third person addressees and the first four with first person addressees, but they occur also with the exact opposite meanings, or comprising more than one person. Here, the term IMPERATIVE is restricted to second person addressees, HORTATIVE to first person addressees, and JUSSIVE to third person addressees. COLLECTIVE COMMANDS denote any COMMAND that applies to more than one person, i.e. any combination of first, second, and third person. PROHIBITIVE refers to a negative COMMAND (‘Don’t do X!’), and POLITE COMMAND to a soft or weak command. There is no number distinction in any of those categories, i.e. no singular or plural distinction. POLITE COMMANDS and PROHIBITIVES can refer to any person addressee. The following section outlines the individual COMMAND types with examples.

The IMPERATIVE is undisputably the canonical COMMAND as found in the literature (e.g. in Aikhenvald 2010). The addressee is a second person, singular or plural, and there is a strong force behind the COMMAND. The speaker directly addresses the intended person and leaves no doubt about his intention; he usually has some kind of authority over the addressee. A straightforward direct second person COMMAND is found in Desano:

(5.67) Desano (Tucanoan; Miller 1999, 73)

yi-re karta goha-beo-ke
 1SG-SPC letter write-send-IMP
 ‘Do write and send me a letter!’

A HORTATIVE has identical speaker(s) and addressee(s). The HORTATIVE suffix *-t^hahta/ -hatahta* in Mamaindê includes the speaker and addressee, but the speaker can also direct the command solely to himself with no other person involved, as in Mapuche. Typically, the latter is translated as ‘Let me...’ and the former with ‘Let us...’.

(5.68) Mamaindê (Nambikwaran; Eberhard 2009, 479)

t^hehati-ijah-tu ?gi-hatahta-wa
 there-DEM-FNS go-HORT-NINTRG
 ‘Let’s go over there!’

(5.69) Mapuche (Araucanian; Smeets 2008, 185)

entu-chi *tífá*
 take.out-HORT1SG this
 ‘Let me take this out.’

The HORTATIVE in Mapuche refers to a proposition or wish (Smeets 2008, 185) and it is arguable whether HORTATIVE has as much force as an IMPERATIVE or rather refers more to a request or suggestion, which would still be a directive, but not a COMMAND.

The JUSSIVE is similarly an arguably low-force directive; it has the least direct connection to the speaker. The addressee is not directly spoken to, and even does not need to be present. The COMMAND takes a detour via a second person that is ordered to do something to a third person, but without an IMPERATIVE marker. This often appears to be a wish or suggestion, rather than command.

(5.70) Hup (Nadahup; Epps 2008, 634)

hūh-way-nih-yi? *nih,* *tih=hup* *tih* *way-ʔúh*
 carry-go.out-NEG-TEL be.IMP 3SG=REFL 3SG go.out-JUS
 ‘Don’t carry him out (of the house); let him go out by himself.’

The COMMAND type with the least force behind it in this sample is the POLITE COMMAND. It may be chosen by a speaker in a defined social setting (e.g. the speaker lacks the sufficient authority over the addressee when the addressee is of higher status), or as a sign of respect or familiarity. In Awa Pit, the relationship of speaker and addressee is not the crucial factor for choosing the POLITE COMMAND suffix *-n(a)ka*, but rather the message content, which borders on a suggestion (Curnow 1997, 246). Whatever the reason, a POLITE COMMAND form implies that the command given is weaker, or softer, than an IMPERATIVE, and closely borders on a request or even a question, but it is not an interrogative as it occurs in non-interrogative clauses. The speaker does expect the command to be carried out, which differentiates it from a real question.

(5.71) Awa Pit (Barbacoan; Curnow 1997, 246)

ayna-t *kwa-nka!*
 cook-SV eat-POL.COM
 ‘Cook and eat it!’ (when giving a present of food)

In Imbabura Quechua, the honorific marker *-pa* can co-occur with the IMPERATIVE marker “referring to the actions of older or respected individuals, or individuals for whom the speaker feels affection” (Cole 1985, 31). Because this is a general politeness marker which also occurs in non-IMPERATIVE clauses, this is not a POLITENESS COMMAND marker.

COLLECTIVE COMMANDS are markers that can refer to any combination of first, second, or third person addressee. These differ from IMPERATIVE, HORTATIVE, and JUSSIVE as those refer to

one specific person only. In Kwaza, the COLLECTIVE COMMAND suffix *-ni* can refer to either first person inclusive or third person:

(5.72) Kwaza (unclassified; Van der Voort 2004, 308)

- (a) *hã 'kui-a-ni*
water drink-1PL-COL.COM
'Let us drink!'
- (b) *ka'we kui-'ni*
coffee drink-COL.COM
'Let him have coffee!'

PROHIBITIVE is any kind of negative COMMAND with the meaning of "Don't do X". This is usually the negation of the IMPERATIVE, but may also include shadings in meaning like threats or warnings (also called admonitives), which mean "Don't do X so that not/because/in order to avoid ...". One and the same marker may be used for PROHIBITIVE and its various semantic offsprings, or each may have its own marker. A typical example for a marker with prototypical PROHIBITIVE meaning is found in Awa Pit (*mun*):

(5.73) Awa Pit (Barbacoan; Curnow 1997, 247)

- kutku-mun!*
tell.a.lie-PROH.SG
'Don't lie!'

Table 5.5 summarizes the correlation of COMMAND type denominations and person and positive/negative, and gives appropriate English examples.

Table 5.5: COMMAND types

COMMAND types/ person	Positive				Negative	Example
	1st person	2nd person	3rd person	More than one person		
HORTATIVE	x	-	-	-	-	Let me go!
IMPERATIVE	-	x	-	-	-	Go!
JUSSIVE	-	-	x	-	-	Let her go!
COLLECTIVE COMMAND	-	-	-	x	-	(All or a selection of the three above)
POLITE COMMAND	x	x	x	x	-	Go, please!
PROHIBITIVE	x	x	x	x	x	Do not go!

5.9.2 Distribution in the sample

At first, the great number of languages marking at least one COMMAND type morpho-syntactically stands out. 52 out of 63 languages in the sample have an IMPERATIVE, HORTATIVE,

JUSSIVE, COLLECTIVE COMMAND, POLITE COMMAND, OR PROHIBITIVE marker. Of the nine languages not marking any COMMAND type, it is possible that COMMAND markers exist in them as well which are just not revealed by the (yet) unsatisfying state of the respective documentation. A candidate is of course Panare, although the documentation for at least Wari', Mocoví, and Pilagá is satisfactory enough to state with modest certainty that there is no morpho-syntactic COMMAND type marking.

JUSSIVES and POLITE COMMANDS are least marked (both 11), followed by COLLECTIVE IMPERATIVES (16) and HORTATIVES (23), and the most marked COMMANDS are PROHIBITIVES (30) and IMPERATIVES (45). Regardless of how many different markers a language may have for one and the same COMMAND type and regardless of which COMMAND type is marked, nine languages do not mark any COMMAND type, 12 mark one, 16 mark two, 16 mark three, seven mark four, two mark five, and one marks all six COMMAND types (see table 5.6).

Formal marking is diverse and ranges from affixes and clitics over particles to auxiliaries. The position of affixes and clitics, however, is surprisingly often to the left of the verb compared to Tense, Aspect, and Modality (other than IMPERATIVE) marking. Still, Bybee's (1985, 173, original highlighting) result that "it is much more common to have an imperative *suffix* than a *prefix*" is confirmed.

A common strategy of expressing COMMANDS is having the bare verbal stem, either with or without person pronouns and with no other TAME marking. Even languages in the sample that do have morpho-syntactic COMMAND marking allow for alternative strategies, such as employing FUTURE or SUBJUNCTIVE. Sometimes these have the same meaning as IMPERATIVE, sometimes they differ in force, e.g. politeness, from the marked IMPERATIVES. For a study of these and command strategies in general the interested reader is referred to the monograph by Aikhenvald (2010). The present study focuses on morpho-syntactic COMMAND marking only, but I wanted to point out that this does not mean that the languages in the sample lack various other means of expressing commands in all their varieties.

The following sections discuss the distribution of COMMAND marking in the sample and comment upon selected peculiarities. The geographical/ genealogical distribution of marking is discussed in section 5.10.

Generalizations

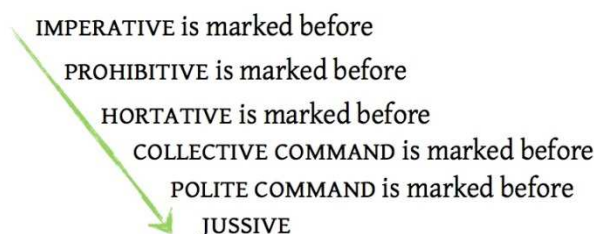
The following generalizations can be made for languages that mark at least one COMMAND type:

1. When a SAIL morpho-syntactically marks a COMMAND type, it is most likely the IMPERATIVE. Languages that do not follow this pattern are Nheengatú, Cubeo, Itonama, Munichí, Yanasha', Paresi, Karo, Mekens, and Kamaiurá. If one includes PROHIBITIVE it is only Cubeo and Itonama that do not fit this pattern. These two languages appear as exceptions throughout the COMMAND type data.
2. When PROHIBITIVE is not marked, IMPERATIVE is always marked (again except for Itonama and Cubeo).

3. When **IMPERATIVE** is not marked, **PROHIBITIVE** is always marked (again except for Itonama and Cubeo).
4. When **HORTATIVE** is marked, **IMPERATIVE** is always marked too.
5. **JUSSIVE** is never the only **COMMAND** marked; at least **IMPERATIVE** is always marked too.
6. When both **JUSSIVE** and **IMPERATIVE** are marked, **HORTATIVE** is marked as well (except for Hup).
7. Polite **COMMANDS** are only marked when at least **IMPERATIVE** is marked too (with the exception of Paresi).

It is possible to generalize the following direction of marking (see also table 5.6):

Figure 5.3: hierarchy of **COMMAND** marking



Given the fact that Cubeo and Itonama do not fit the general pattern of **COMMAND** marking, they will now be looked at in detail.

In Cubeo, the suffix **-xA** is found in **COMMANDS** referring to all three persons. The verb receives the **COLLECTIVE COMMAND** suffix and the respective person pronouns. The first person plural **HORTATIVE** has an alternative construction without **-xA** which implies more urgency (Morse & Maxwell 1999, 24-26). No other **COMMAND** marker could be found in Morse and Maxwell (1999). The other Tucanoan language in the sample is Desano, which shows a completely different **COMMAND** marking system (**IMPERATIVE**, **HORTATIVE**, **JUSSIVE**). This is not surprising, because, according to Barnes (1999, 209), they belong to quite different branches: Cubeo is Central Tucanoan and Desano Eastern Central Tucanoan. Another Tucanoan language not in the sample, Tucano (Eastern Northern Tucanoan according to Ethnologue), has an **IMPERATIVE** suffix (**-ya/ -ña/ -a**), but no **COLLECTIVE COMMAND** either (West 1980, 48ff.). Further research is necessary to establish whether Cubeo is an exception in **COMMAND** marking not only in the present sample, but also within the Tucanoan family.

The prefix that marks **IMPERATIVE**, **HORTATIVE**, and **JUSSIVE** in Itonama (**ki-**) is homophonous with the **INTENTIONAL** prefix. The relationship between intention and command is easy to see, but the direct relations are yet to be investigated. Certainly, this prefix is not a typical **COMMAND** marker and this may explain why Itonama does not fit the generalization pattern above. Until more data becomes available, **ki-** is treated as ambiguous morpheme (between **COMMAND** and **INTENTIONAL**).

(5.74) Itonama (unclassified; Crevels p.c.)

- (a) *dih-ki-yaka-ne* *waripahoro*
 1PL.INCL-IMP/HORT/JUS/INT-sing-N *waripahoro*
 ‘Let’s sing the *waripahoro*.’
- (b) *a’-ki-yaka-ne*
 2SG-IMP/HORT/JUS/INT-sing-N
 ‘Sing!’
- (c) *ko-pehkele’-na* *se’-pa-chihî-ke*
 IMP/HORT/JUS/INT-be.quiet-N 1SG.INV-speech-bother-PL
 ‘Let him be quiet, he bothers me with his noise.’
- (d) *chaphoko* *mama’na* *se’-payk’i-ne* *assésiyo* *ni-k’a’ne*
 tomorrow AUX:FUT 1SG.INV-bring-N Ascensio CLF:sinuous-one
pak’isine *as-mi-kareta* *as-ki-nos-kos-ch’i’-ne*
 CONJ:PURP 1SG-REL-oxcart 1SG-IMP/HORT/JUS/INT-up-side-attach-N
 ‘Tomorrow Ascensio is going to bring me a yoke of oxen so that I can attach my
 oxcart to the yoke.’

While it is not possible to discuss every detail about SAILS’ COMMAND marking, a few topics of interest that emerge out of the data are now highlighted: (i) three languages that exhibit FAMILIAR COMMAND markers, (ii) MOTION COMMANDS (space and time).

Three languages have distinct markers for COMMANDS that apply to family or friends, or in general familiar persons, and which are, for lack of a better term, called here FAMILIAR COMMAND: Aguaruna, Yaminahua, and Mamaindê. They are probably not simply subtypes of POLITE COMMANDS, because the latter two possess a different marker for POLITE COMMANDS (although it is of course possible that POLITE COMMAND markers in other languages in the sample include the possibility to use them with familiars), and because the Mamaindê marker clearly has more force than either IMPERATIVE OR POLITE COMMAND (see below).

In Aguaruna, the FAMILIAR COMMAND suffix *-kia* is used in COMMANDS that are directed towards family members or friends (Overall 2007, 351). A POLITE COMMAND is expressed by a combination of the IMPERATIVE suffix *-ta* and the attenuative suffix *-sa* (ibid. 292). The IMPERATIVE is clearly the strongest COMMAND of the three, but the exact difference in meaning between a FAMILIAR and a POLITE COMMAND remains to be investigated.

Unlike Aguaruna, Yaminahua does have a separate POLITE COMMAND marker (suffix *-kî*) which is a softer COMMAND than IMPERATIVE (suffix *-fe*), but again, the exact difference between POLITE and FAMILIAR COMMAND is unknown (except for the fact that the familiar COMMAND often co-occurs with the urgency suffix *-xto*) (Faust & Loos 2002, 159f.).

In Mamaindê, the normal IMPERATIVE suffix *-tahî* is contrasted with the POLITE COMMAND suffix *-tsîn*. Additionally, the FAMILIAR COMMAND suffix *-tah* is used in informal and colloquial contexts and expresses the most force (Eberhard 2009, 478f.). This marker is certainly completely different from those in Aguaruna and Yanasha’, although directed to roughly the same audience. But whereas the increased level of familiarity and intimacy between

speaker and addressee leads to an increased level of politeness in the first two languages, in Mamaindê it leads to an almost pejorative usage.

The following examples first give the IMPERATIVE, then the POLITE COMMAND or alternative construction, and then the FAMILIAR COMMAND markers.

(5.75) Aguaruna (Jivaroan; Overall 2007, 351, 292, 351)

- | | | |
|-----|--|--------------------------------|
| (a) | <i>mai-ta</i>
bathe+LOAF-IMP
'Have a bath!' | IMPERATIVE |
| (b) | <i>uwa-sa-ta</i>
drink-ATT-IMP
'Drink!' | POLITE COMMAND
construction |
| (c) | <i>mai-kia</i>
bathe+LOAF-FAM.IMP
'Have a bath!' | FAMILIAR COMMAND |

(5.76) Yaminahua (Panoan; Faust and Loos 2002, 159, 160)

- | | | |
|-----|--|------------------|
| (a) | <i>kirika ane-fe</i>
book read-IMP
'Read the book!' (orig. 'Lee el libro.') | IMPERATIVE |
| (b) | <i>ẽ mia yoi-nõ, nika-kĩ</i>
I you talk-COL.COM listen-POL.COM
'Let me talk to you, listen.' (orig. 'Te voy a hablar, escucha.') | POLITE COMMAND |
| (c) | <i>o-pe</i>
come-FAM.IMP
'Come!' (orig. 'Ven.') | FAMILIAR COMMAND |

(5.77) Mamaindê (Nambikwaran; Eberhard 2009, 479, 345, 479)

- | | | |
|-----|--|------------------|
| (a) | <i>?ain-ã sun-tahĩn-wa</i>
fish-FNS kill-IMP-N.INT
'Kill the fish!' | IMPERATIVE |
| (b) | <i>ta?lohna-ki? tanu-tsĩn-wa.</i>
old.woman-incl-also give-POL.COM-N.INT
'Give (some) to the old woman as well.' | POLITE COMMAND |
| (c) | <i>ha?tin ana?-tah</i>
quickly stop-FAM.COM
'Stop quickly!' | FAMILIAR COMMAND |

POLITE COMMANDS are found in many languages as a by-product of Modality marking different from COMMAND marking. For example, in Wari' and Baure, the IRREALIS markers can refer to a POLITE COMMAND and requests and suggestions.²⁴

MOTION COMMANDS specify motion, either to or away from the speaker and in general involve a motion verb. A MOTION COMMAND away from the speaker is sometimes referred to as 'venitive' ('Come and do X!'), towards the speaker as 'allative' (Go and do X!).²⁵ Because a few languages in the SAILS sample do exhibit MOTION COMMANDS, they will be now discussed.

Both Cariban languages in the sample (Hixkaryana and Tiriyo) make a distinction of NONMOTION VS. MOTION COMMAND marking. Hixkaryana has two different paradigms for commands involving motion and nonmotion. The nonmotion paradigm has cumulative morphemes for the full range of persons (1st, 2nd, 1stincl, 3rd) and number (individual/collective), but the motion paradigm only for 1st and 2nd person and number (Derbyshire 1985, 194f.). Hixkaryana MOTION COMMANDS are used when the verb in the COMMAND clause is a motion verb and seem restricted to allative meanings.

Tiriyo has a similar system: IMPERATIVES that do not involve motion away from the speaker are marked by *-kë*, whereas IMPERATIVES that do are marked by *-ta* (Meira 1999, 319) (compare with the MOTION IMPERATIVE form *-ta* in Hixkaryana).

(5.78) Hixkaryana (Cariban; Derbyshire 1985, 195, 195)

- (a) *a-wanota-ko*
2SG-sing-NMOT.IMP
'Sing!'
- (b) *Ø-ewehi-ta*
GNRF-take.a.bath-MOT.IMP
'Go take a bath.'

(5.79) Tiriyo (Cariban; Meira 1999, 319, 319)

- (a) *tíwaarë eh-kë*
careful COP-NMOT.IMP
'Be careful!'
- (b) *papa_pa i-wa-t*
2:father_REP 3O-fetch-MOT.IMP
'Go get your father!'

In Cocama-Cocamilla, a venitive MOTION COMMAND is marked by the construction verb + venitive auxiliary *uri* and subject suppression or the particle *tsaniuri*, and an allative MOTION

²⁴ Interestingly, the FRUSTRATIVE marker =*chi* in Yurakaré has politeness value in combination with COMMANDS (Gipper 2011, 187).

²⁵ SAILS not in the sample which have MOTION COMMANDS include Wanano and Barasano (both East Tucanoan) (Aikhenvald 2010, 374ff.).

COMMAND by the particle *yawa* (Vallejos Yopán 2010, 548ff.). The canonical IMPERATIVE without venitive meaning is unmarked and expressed by the bare verb with a special intonation pattern. The form *uri*, which exists as the independent verb ‘come’, also occurs in the POLITE IMPERATIVE morpheme *tsaniuri* (with *tsani* ‘try’) and expresses an invitation to come and do X; the auxiliary *uri* can co-occur with this form but this is not necessary. Ika has a venitive MOTION COMMAND that is similar in form to the particle *yawa*: the auxiliary *awa*.²⁶

(5.80) Cocama-Cocamilla (Tupían, Guaraní; Vallejos Yopán 2010, 548, 548, 555, 560)

- (a) *utsu*
go
‘Go!’
- (b) *ey-uri*
eat-AUX
‘Come to eat!’
- (c) *yawa yatsuka*
MOT.IMP take.bath
‘Go take a bath!’
- (d) ***tsaniuri*** *yatsuk=uri*
MOT.POL.IMP take.bath=AUX
‘Come and take a bath.’

(5.81) Ika (Chibchan, Aruak; Frank 1985, 118)

- awakati dže ido-awa*
avocado water spill-AUX
‘(Go) water the avocado tree! (over there/ away)’

In addition to COMMANDS that are near or removed in space, COMMANDS can be marked according to whether they are near or removed in time. Tariana has a special marker *-wa* for a postponed COMMAND (Aikhenvald 2003, 373). Jarawara partially conflates these meanings: it has IMPERATIVE markers for ‘here and now’ and ‘in some distant time or place’ (Dixon 2004, 397). In Tiriyó, the IRREALIS marker co-occurring with IMPERATIVE marking indicates that the action is to be carried out later.

²⁶ The element *awa* occurs as a noun in Cocama-Cocamilla meaning ‘people, human being’ and as interrogative ‘who’, in the verb *yawachima* ‘arrive’, and in the adverb *kawa* ‘far away’ (found in Vallejos Yopán 2010). From these, the latter form is possibly related to the MOTION COMMAND.

(5.82) Tariana (Arawakan; Aikhenvald 2003, 373)
diwese-wya pi-a-wa nuitō
 then-EXTRAL 2SG-go-IMP 1sg.daughter:VOC
 ‘Go a little later, my daughter!’

(5.83) Jarawara (Arawakan; Dixon 2004, 397, 379)

(a) *otara noki ti-na-hi*
 1EXCL.OBJ wait 2SG.SBJ-AUX-IMM.IMP.F
 ‘You (sg) wait for us (here and now)!’

(b) *otara noki ti-jahi*
 1EXCL.OBJ wait 2SG.SBJ-DIST.IMP.F
 ‘You (sg) wait for us (in some distant place or time)!’

Figure 5.1: Number of SAILs marking COMMAND types



Figure 5.2: Number of SAILs marking number of COMMAND types

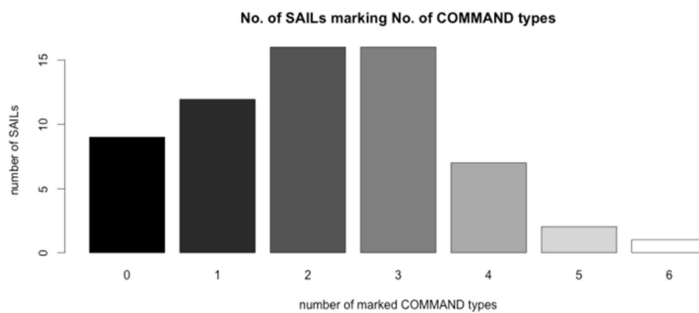


Table 5.6: COMMAND type marking in the SALLs

	LANGUAGE	FAMILY	IMP	PROH	HORT	COL COM	POL COM	JUS	TOTAL
1	Panare	Cariban							0
2	Wari'	Chapacuran							0
3	Chimila	Chibchan							0
4	Tehuelche	Chonan							0
5	Mocoví	Guaycuruan							0
6	Pilagá	Guaycuruan							0
7	Bororo	Macro-Gêan							0
8	Kaingang	Macro-Gêan							0
9	Timbira	Macro-Gêan							0
10	Tsafiki	Barbacoan							1
11	Mosetén	Mosetenan							1
12	Embera	Chocoan							1
13	Cuzco Quechua	Quechuan							1
14	Kanoê	Unclassified							1
15	Karitiana	Tupían							1
16	Warao	Unclassified							1
17	Nheengatú	Tupían, Guaraní							1
18	Yanesha'	Arawakan							1
19	Munichi	Unclassified							1
20	Itonama	Unclassified							1
21	Cubeo	Tucanoan							1
22	Jarawara	Arawan							2
23	Dâw	Nadahup							2
24	Matses	Panoan							2
25	Trumai	Unclassified							2
26	Sateré-Mawé	Tupían							2
27	Apurinã	Arawakan							2
28	Hixkaryana	Cariban							2
29	Rikbaktsa	Macro-Gêan							2
30	Sabanê	Nambikwaran							2
31	Miraña	Boran							2
32	Shipibo- Konibo	Panoan							2
33	Hup	Nadahup							2

34	Kamaiurá	Tupían, Guaraní							2
35	Karo	Tupían							2
36	Mekens	Tupían							2
37	Paresi	Arawakan							2
38	Ika	Chibchan, Aruak							3
39	Nasa Yuwe	Paezan							3
40	Wichí (Mataco)	Matacoan							3
41	Yanam	Yanomaman							3
42	Emérillon	Tupían, Guaraní							3
43	Tapiete	Tupían, Guaraní							3
44	Yurakaré	Unclassified							3
45	Puinave	Unclassified							3
46	Awa Pit	Barbacoan							3
47	Movima	Unclassified							3
48	Tiriyó	Cariban							3
49	Mamaindê	Nambikwaran							3
50	Baure	Arawakan							3
51	Huallaga Quechua	Quechuan							3
52	Imbabura Quechua	Quechuan							3
53	Desano	Tucanoan							3
54	Kwaza	Unclassified							4
55	Cocama- Cocamilla	Tupían, Guaraní							4
56	Mapuche	Araucanian							4
57	Urarina	Unclassified							4
58	Yaminahua	Panoan							4
59	Aymara	Aymaran							4
60	Leko	Unclassified							4
61	Tariana	Arawakan							5
62	Aguaruna	Jivaroan							5
63	Cavineña	Tacanan							6
			45	30	23	16	11	11	136/ 378

5.10 TYPOLOGICAL DISTRIBUTION IN THE SAMPLE

I will begin by consolidating the findings from sections 5.2 to 5.9, aiming to give an overview of Modality marking in the SAILS in the sample. Because individual results were discussed in the respective sections, they will not be repeated here. Instead, I present the distribution of all Modality markers and their typological implications and discuss the ramifications for the study of grammaticalization.

In the sample, 39 languages mark *DESIDERATIVE*, 28 mark *DUBITATIVE* and *IRREALIS*, 24 mark *PURPOSIVE*, 23 mark *FRUSTRATIVE*, 20 mark *CERTAINTY*, 17 mark *POTENTIAL*, 13 mark *INTENTIONAL*, and three mark *REALIS*. Altogether 52 out of 63 languages mark at least one *COMMAND* type. When the individual *COMMAND* types are taken into consideration, the following hierarchy emerges: *IMPERATIVE* (45), *DESIDERATIVE* (39), *PROHIBITIVE* (30), *DUBITATIVE* (28) and *IRREALIS* (28), *PURPOSIVE* (24), *FRUSTRATIVE* (23) and *HORTATIVE* (23), *CERTAINTY* (20), *POTENTIAL* (17), *COLLECTIVE COMMAND* (16), *INTENTIONAL* (13), *POLITE COMMAND* (11) and *JUSSIVE* (11), and *REALIS* (3).

No language marks all nine Modalities, but only one marks no Modality (Chimila, which does not mark any *COMMAND* type either). Nine languages mark one Modality, 18 mark two, nine mark three, 11 mark four and five, three mark six and one marks seven. Figures 5.4 and 5.5 plot the number of languages marked per category and the number of languages marking the number of Modals (from zero to seven), respectively. It can be seen that *COMMAND* marking (all *COMMAND* types collapsed) occurs most frequently and that *REALIS* is least marked, and that languages marking two Modal categories are most frequent in contrast to those marking one and seven categories. Table 5.7 shows which language exhibits marker for which category, ordered from least marking to most marking (this table excludes *COMMAND* types which are presented in table 5.6).

In contrast to the *COMMAND* types systems (cf. section 5.9), there are very few implications or relations in the Modality sample excluding the *COMMAND* types. Modality marking seems chaotic throughout the sample, although the *COMMAND* types distribution by itself performs better than the other nine Modalities (see below).

Previous research has suggested that Agent-oriented Modality is rarely marked inflectionally cross-linguistically, i.e. by non-bound morphemes, and that the other Modalities tend to be more bound in comparison (Bybee 1985, 166; Bybee et al. 1994, 241). This cannot be confirmed completely for the present study. Because this study does not code for periphrastic constructions there can be no comparison with periphrastic marking as in Bybee et al. (1994, 242). I will now comment upon the boundedness of forms in the present Modality sample.

The two Agent-oriented Modalities in the present sample, *INTENTIONAL* and *DESIDERATIVE*, pattern differently on the Modality scale. Whereas *INTENTIONAL* indeed ranges quite low (13), *DESIDERATIVE* is the second-highest marked Modality after *COMMAND* (36). Of the 39 languages marking *DESIDERATIVE*, 33 have suffixation, two have auxiliaries, four have particles, and one has as marker that may be either a clitic or particle (Bororo). When a language has more

than one *DESIDERATIVE* marker, they always have the same form; e.g. both *DESIDERATIVE* markers in Sabanê are suffixes.²⁷ Of the 13 languages marking *INTENTIONAL*, 10 have suffixation, two have particles, one has a prefix, and one has a clitic (Baure marks *INTENTIONAL* with a particle and a suffix). Even if one considers particles and auxiliaries as non-bound, the overwhelming majority of *DESIDERATIVE* and *INTENTIONAL* marking languages still has bound markers: *DESIDERATIVE*: in 33 out of 39 cases, *INTENTIONAL* in 12 out of 13. This is not surprising, because “the tendency of a gram to affix to a stem is partly governed by the semantic relevance of that gram to the stem” (Bybee et al. 1994, 242). *DESIDERATIVE* and *INTENTIONAL* markers directly modify the verb and are thus highly relevant to the verbal semantics, and more likely to be bound. However, this needs further verification in comparison to non-morpho-syntactical ways of expressing *DESIDERATIVE* and *INTENTIONAL* in the same set of languages.

Bybee et al. (1994, 241) also report that “imperatives are most likely to be bound”. The *IMPERATIVE* certainly is the category in the sample that is marked most (in 45 languages), and like the *DESIDERATIVE* and *INTENTIONAL*, it is mostly marked by suffixing. If we assume that categories with a high amount of bound morpho-syntactic markers (compared to the amount of non-bound forms) are more likely to be at a later stage of grammaticalization (as suggested by Bybee et al. 1994), this leads to the hypothesis that all Modal categories in the sample are highly grammaticalized. This certainly holds true for *INTENTIONAL* and *DESIDERATIVE*, although Bybee et al. (1997, 241) argue that “agent-oriented modalities generally are at early stages of grammaticalization, while the other modalities are generally at later stages”. Early stages of grammaticalization involving periphrastic expressions are not taken into account here. But because this study does consider both bound (e.g. affixes and clitics) and non-bound (e.g. particles and auxiliaries) markers, it is possible to establish the relative position of a marker on a grammaticalization path, and subsequently of Modal categories in a cross-linguistic perspective as well. It is however debatable whether boundedness in this case equals a high degree of grammaticalization. In general, SAILS have been shown to possess a high degree of complex morphology not only for *TAME* but various other functions. This may mean that all these functions have a high degree of grammaticalization, or, rather more probable, that boundedness does not automatically entail a high degree of grammaticalization.

At this point, a note about the sources of Modality markers in the sample is in order. The origins of *DESIDERATIVE* markers are discussed in Mueller (2013). They often originate in words with the meaning ‘want’ or *FUTURE* markers, or that both *FUTURE* and *DESIDERATIVE* have a common source. *FUTURE* and Modal categories are in general strongly related in terms of common origins in the sample, as demonstrated for the case of Aguaruna below.

The striking similarity of the *FUTURE*, *IMPERATIVE*, *JUSSIVE*, *INTENTIONAL*, *FRUSTRATIVE*, and *DESIDERATIVE* markers in Aguaruna (all involve /ta/) suggests a common origin. According to

²⁷ Hixkaryana is the only exception with a suffix and a particle (postposition).

Overall (2007, 347) ‘desirable’ FUTURE was the original meaning and the other Modals developed from it (see also below). Epps (2008, 640) argues for a grammaticalization path of the DUBITATIVE marker in Hup from a directive to an Epistemic marker, or, in Palmer’s terms, from Deontic to Epistemic, which is not uncommon (cf. Palmer 2001, 87-89). Bybee et al. (1994, 195) demonstrate as well that Epistemic Modals often develop from Agent-oriented Modality.

The CERTAINTY particle *llemay* in Mapuche consists of the affirmative particle *lle* and the particle *may* which is used in questions to “provoke an affirmative answer” (Smeets 2008, 334), and the DUBITATIVE particle *ni* in Cavineña is probably related to the negator *ni*, which is borrowed from Spanish *ni* ‘not even’ (Guillaume 2008, 642). Nheengatú borrowed the phrase *será que* as DUBITATIVE particle *seraki* from Portuguese (da Cruz 2011, 361). All in all, not much is known about the sources of most markers, but where we do know these they point towards origins of lexemes grammatical markers with related meanings. One could infer the origin of a language-specific marker by its current grammaticalized form and parallel, established grammaticalization paths. In cases where no origin can be found within a language or language family that would suggest possible borrowing. Thus, backtracking grammaticalization paths provides a useful tool for uncovering language contact situations.

Based on the assumption that the Tariana IMPERATIVE marker *-sí* is a reinterpretation of the FUTURE marker *-sí* (Aikhenvald 2010, 381),²⁸ I did a survey of correlations of forms between FUTURE and related categories (e.g. INTENTIONAL, PURPOSIVE, IRREALIS) and COMMAND type markers. COMMANDS and FUTURE Tense are naturally related, because prototypically a command refers to an action in the future (cf. Aikhenvald 2010, 128-133). Bybee et al. (1994, 273) report that 13 languages in their sample use FUTURE markers to express COMMANDS, and postulate that COMMAND markers grammaticalize from FUTURE markers.

In general, however, the forms in the present sample do not show overt formal correlations, with a few exceptions: homophonous markers occur in Aguaruna (FUTURE, IMPERATIVE, JUSSIVE), Itonama (INTENTIONAL, COLLECTIVE COMMAND). Aguaruna employs the suffix *-ta* for FUTURE, IMPERATIVE and JUSSIVE. The same form occurs in the DESIDERATIVE, FRUSTRATIVE, and INTENTIONAL markers. As noted, Overall (2007, 347) suggests that ‘desirable’ FUTURE was probably the original meaning and that the other Modal shades have developed from this one marker. This confirms the grammaticalization path of Bybee et al. (1994).

(5.84) Aguaruna (Overall 2007, 351, 23, 354, 303)

- | | | |
|-----|----------------|------------|
| (a) | <i>mai-ta</i> | IMPERATIVE |
| | bathe+LOAF-IMP | |
| | ‘Have a bath!’ | |

²⁸ It is not clear on what basis Aikhenvald (2010) defines *-sí* as FUTURE marker. The FUTURE markers in Tariana in Aikhenvald (2003, 320) are *-mhade* and *-de*.

- | | | |
|-----|---|--------------|
| (b) | <i>yu-a-ta-ha-i</i>
eat-HIAF-FUT-1SG-DECL
'I will eat this.' | FUTURE |
| (c) | <i>yu-a-tata-ha-i</i>
eat-HIAF-DES-1SG-DECL
'I want to eat.' | DESIDERATIVE |
| (d) | <i>impi-mitika-takama</i>
become.deaf-CAUS-FRUST
'trying in vain to deafen (him)...'
 | FRUSTRATIVE |
| (e) | <i>mai-tasa-nu</i> <i>wi-a-ha-i</i>
bathe:LOAF-INT-1SG:SS go-IPFV-1SG-DECL
'I'm going to bathe.' | INTENTIONAL |

A rather unusual correlation is found in Matses, where the suffix *-enda* occurs as the FUTURE as well as the PROHIBITIVE morpheme. It does not present a problem to connect *-enda* with Modal values, but it is unclear how it acquired a negative meaning, if it really is the same marker. Usually, FUTURE in Matses is marked by the NONPAST suffix *-e*, and when *-enda* is used there are Modal overtones specifically of politeness, requests, and uncertainty (ibid. 428-429). It attaches directly to the verb stem and is followed by person and/ or Mood marking. When a verb is marked with PROHIBITIVE *-enda* it does not receive any further marking. The PROHIBITIVE *-enda* possibly consists of the negation marker *-en* and IMPERATIVE *-ta*, but Fleck (2003, 444) rejects this on phonological and syntactic grounds. Nevertheless, the grammaticalization path of the PROHIBITIVE marker *-enda* in Matses does seem to involve IMPERATIVE and FUTURE meanings, although in which way exactly has to be further researched.

Incidentally, Matses marks IMPERATIVE by *-ta*, just like Aguaruna. That marker apparently has a Jivaroan background, because it is not only found in Aguaruna, but in Shuar as well: HORTATIVE *-ta* and FUTURE *-tta* (Adelaar with Muysken 2004, 441-442). In the two Panoan languages in the sample, Yaminahua²⁹ and Shipibo-Konibo, no similar form could be found either as FUTURE or COMMAND type.

(5.85) Matses (Panoan; Fleck 2003, 444, 429, 428)

- | | |
|-----|--|
| (a) | <i>pos-enda</i>
shatter-PROH
'Don't break it!' |
| (b) | <i>mibi nid-enda-c</i>
2ABS go-NPST-IND
'You might go.' |

²⁹ Yaminahua has a suffix *-ta* for either distal or "PAST of yesterday" but there is no relationship with FUTURE or COMMAND types.

- (c) *në-bi* *nid-enda-bi*
 now-EMP go-NPST-1SG
 'I'm leaving now (so hurry up if you don't want to be left behind).'

Another correlation between unrelated languages of COMMAND and FUTURE forms occurs with the suffix *-ya* as COMMAND in Tariana (Arawakan), Barasano, Tuyuca, Desano (all three Tucanoan), and as FUTURE in Embera (Chocoan) and Puinave (unclassified). These languages partially coincide with the proposed linguistic are Içana-Vaupés (cf. Aikhenvald 1999) and adjoining areas, but more research is necessary to establish a certain contact phenomenon here.

There are several cases of identical forms throughout the sample and Modality/ Tense categories that occur in unrelated and/or geographically distanced languages. These are interesting in the light of possible contact phenomena, but it is often not possible to establish a geographical or cultural link between them. Most likely the similarities are due to chance. They will be briefly listed below, in the hope that future research will reveal their origins. The best case for a possible contact phenomenon is the existence of IMPERATIVE *-ma* and JUSSIVE *-pa* in Aymara, and IMPERATIVE singular *-ma* and IMPERATIVE plural *-pa* in Yurakaré. Less plausible is the relationship between the IMPERATIVE prefixes *m-* in Nasa Yuwe and Yurakaré; the PROHIBITIVE suffixes *-ume* in Kamaiurá and Cavineña; the PURPOSIVE suffixes *-ra* in Cocama-Cocamilla and Cavineña; the IMPERATIVE suffixes *-wa* in Trumai, Tariana, and Mosestén (/wa/ occurs in Cocama-Cocamilla IMPERATIVE particle *yawa*); or the element /ki/ as found in IMPERATIVE forms in Aguaruna, Yanam, Movima, and Trumai (and in the PROHIBITIVE suffix in Mapuche and in the Itonama COLLECTIVE COMMAND).

Figure 5.4: Number of SAILs marking Modality (COMMAND types as one column)

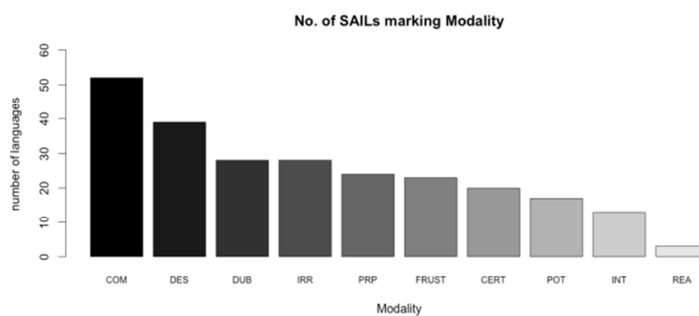


Figure 5.5: Number of SAILS marking number of Modalities

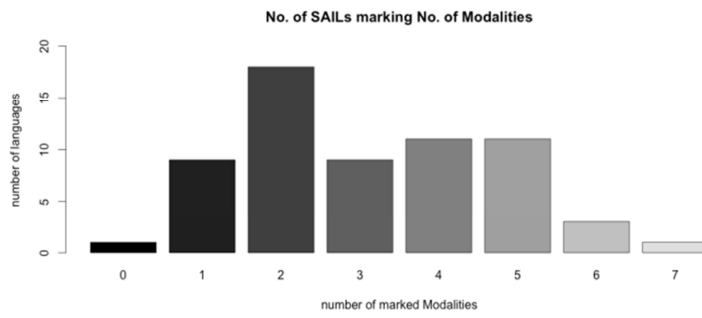


Table 5.7: Modality marking in the SAILs

	LANGUAGE	FAMILY	DES	DUB	IRR	PRP	FRUST	CERT	POT	INT	REA	TOTAL
1	Chimila	Chibchan										0
2	Kanoê	Unclassified	■									1
3	Mocoví	Guaycuruan	■									1
4	Pilagá	Guaycuruan	■									1
5	Sabanê	Nambikwaran	■									1
6	Tapiete	Tupían, Guaraní	■									1
7	Kaingang	Macro-Gêan		■								1
8	Embera	Chocoan				■						1
9	Yanam	Yanomaman							■			1
10	Itonama	Unclassified								■		1
11	Panare	Cariban	■		■							2
12	Huallaga Quechua	Quechuan	■		■							2
13	Karitiana	Tupían	■		■							2
14	Shipibo- Konibo	Panoan	■			■						2
15	Trumai	Unclassified	■			■						2
16	Munichi	Unclassified	■			■						2
17	Aymara	Aymaran	■					■				2
18	Urarina	Unclassified		■	■							2
19	Karo	Tupían		■		■						2
20	Yanasha'	Arawakan		■				■				2
21	Timbira	Macro-Gêan			■		■					2
22	Wari'	Chapacuran			■						■	2
23	Tehuelche	Chonan			■						■	2
24	Wichí (Mataco)	Matacoan			■						■	2
25	Miraña	Boran				■	■					2
26	Ika	Chibchan , Aruak				■				■		2
27	Rikbaktsa	Macro-Gêan					■	■		■		2
28	Nasa Yuwe	Paezan							■	■		2
29	Bororo	Macro-Gêan	■	■	■							3
30	Hup	Nadahup	■	■	■	■	■					3

31	Sateré-Mawé	Tupían											3
32	Leko	Unclassified											3
33	Dâw	Nadahup											3
34	Aguaruna	Jivaroan											3
35	Paresi	Arawakan											3
36	Cuzco Quechua	Quechuan											3
37	Cocama- Cocamilla	Tupían, Guaraní											3
38	Imbabura Quechua	Quechuan											4
39	Emérillon	Tupían, Guaraní											4
40	Cubeo	Tucanoan											4
41	Desano	Tucanoan											4
42	Awa Pit	Barbacoan											4
43	Apurinā	Arawakan											4
44	Jarawara	Arawan											4
45	Mapuche	Araucanian											4
46	Nheengatú	Tupían, Guaraní											4
47	Tariana	Arawakan											4
48	Baure	Arawakan											4
49	Tsafiki	Barbacoan											5
50	Kamaiurá	Tupían, Guaraní											5
51	Hixkaryana	Cariban											5
52	Cavineña	Tacanan											5
53	Yaminahua	Panoan											5
54	Matses	Panoan											5
55	Puinave	Unclassified											5
56	Kwaza	Unclassified											5
57	Mamaidê	Nambikwaran											5
58	Warao	Unclassified											5
59	Mosetén	Mosetenan											5
60	Mekens	Tupían											6
61	Yurakaré	Unclassified											6
62	Movima	Unclassified											6
63	Tiriyó	Cariban											7
			39	28	28	24	23	20	17	13	3		195

5.11 GEOGRAPHICAL DISTRIBUTION

This section comments on geographical patterns of marked Modal categories. The following paragraphs investigate clusters of Modal marking, including *FRUSTRATIVE*, *DESIDERATIVE*, Epistemic Modality, and *COMMAND* types, and discusses links between Modal markers and Tense markers.

The number of Modals occurring in the sample (without *COMMAND* types) marked by languages ranges from zero to seven (maximally possible are nine). Map 5.1 shows the distribution of the number of Modal marking in the sample. While the different parts of the continent vary greatly in how many Modals are marked, the Vaupés area features a cluster of languages marking four Modals (Tariana, Cubeo, Desano, Nheengatú; Dâw has three). Additionally, most languages with the highest number of marked Modals (six and seven) occur in the Guaporé-Mamoré and adjoining regions (Yurakaré, Mekens, Movima). A minor cluster of languages marking two Modals can be seen in central to north Peru (Urarina, Munichí, Huallaga Quechua, Shipibo-Konibo, and Yanasha). The high diversity of Modality marking in the Guaporé-Mamoré area, not only in terms of number of Modals but also in terms of variation of Modal categories, is explained by the high number of unrelated and especially isolates or unclassified languages in that region (cf. Crevels & Van der Voort 2008). There are 17 of the languages proposed for the Guaporé-Mamoré by Crevels & Van der Voort (2008) which are also part of the present sample, including seven unclassified languages or isolates, and Tupían, Tacanan, Arawakan, Panoan, Nambikwaran, Macro-Gêan, and Chapacuran languages. Interestingly, the variation of Modality marking in that area outweighs the similarities, but the limitations of the sample do not allow stating this with certainty.

Contrary to most categories in this sample, *FRUSTRATIVE* exhibits a pattern of marking which is quite frequent in certain regions. The frequent occurrence of *FRUSTRATIVE* in *SAILS* has been noted before, and indeed this was why the feature was chosen for this study. Aikhenvald (1999, 94) mentions that *FRUSTRATIVE* often occurs in Arawakan languages³⁰ and Campbell (2012, 291) adds that “frustrative grammatical markers are not unknown in other parts of the world, but they seem especially well represented in SA in comparison to elsewhere”.³¹ Payne (1990, 216) mentions *FRUSTRATIVE* marked by suffixation in Tupían (Guaraní), as does Dietrich (2006) (see also Overall, 2012).³² *FRUSTRATIVE* marking is a quite exotic feature for speakers of an Indo-European language, which may lead to the suggestion

³⁰ Of the Arawakan languages in the present sample, Paresí, Tariana, and Apurinã mark *FRUSTRATIVE*, Baure and Yanasha’ do not.

³¹ Campbell (2012) lists Urarina and Baure among the languages which have *FRUSTRATIVE*. Urarina does have a marker with *FRUSTRATIVE*-like functions, but it does not fit the present definition; Baure clearly has no *FRUSTRATIVE* marker, although the term appears in historical sources (Danielsen p.c.). Campbell does not discuss his sources here.

³² For a study of the *FRUSTRATIVE* morpheme *-pana* functioning as conjunction in Amahuaca (Panoan) see Sparing-Chavez (2003).

that its occurrence in slightly over a third of the SAILS in the sample is statistically significant. Unfortunately, in the light of the absence of studies outside SA it is not possible to test this. Aikhenvald (2012, 185) suggests that FRUSTRATIVE could have developed into a characteristic of certain areas, such as the Vaupés and Amazonia in general:

“[f]rustrative markers are found in almost every family in Amazonia. Pervasive as they are in subgroups such as Tupí-Guaraní, it appears to be hard to reconstruct a common form. Only some frustrative markers within Panoan languages have cognates across the family. The frustrative is a feature of all the languages in the linguistic area of the Vaupés and its surrounds [...]. The distribution of the frustrative, as a semantic package, could well be a combination of areal diffusion and genetic inheritance”.

Languages of the Vaupés occurring in the sample are Tariana, Cubeo, Desano, Nheengatú, and Dâw. All these have FRUSTRATIVE markers, although these do not seem to be related. The Nheengatú FRUSTRATIVE is a Tupían cognate (see table 5.12), the FRUSTRATIVE auxiliary *wud* in Dâw originates in the verb ‘be, possibility to be, not being full’ (Martins 2004, 293); the other sources are unknown. This points toward independent development in each language although possibly induced by contact. Nheengatú as *lingua franca* offers itself as model language, but this needs further study.

Table 5.8: FRUSTRATIVE markers in the Vaupés

Language	Tariana	Desano	Cubeo	Nheengatú	Dâw
Family	Arawakan	Tucanoan	Tucanoan	Tupían, Guaraní	Nadahup
FRUSTRATIVE	Clitic <i>-tha</i>	Suffix <i>-ri/ -ra</i>	Suffix <i>-du</i>	Particle <i>yepé</i>	aux <i>wud</i>

Dietrich (2006, 71) claims that FRUSTRATIVE is especially prominent in Amazonia which is confirmed in this sample (see map 5.2).

FRUSTRATIVES as occurring in the sample do not suggest common origins within language families, with the exception of Tupían (see section *Tupían* below), where the forms hint towards two different sources. This points towards a combination of diffusion of the FRUSTRATIVE semantic category rather than the form, in addition to common origins in at least Tupían, just as Aikhenvald suggested (although the three Panoan languages in the sample (Shipibo-Konibo, Matses, Yaminahua) do not possess morpho-syntactic FRUSTRATIVE markers).

The same factors have already been argued to be behind the distribution of DESIDERATIVE in SA (Mueller 2013). Although FRUSTRATIVE occurs less often in the sample than DESIDERATIVE, it is of course possible that it spread by the same processes. Are the same languages or areas affected by the developments of both features? FRUSTRATIVE markers in the present sample occur clustering in and around the Vaupés and to a certain extent in Amazonia and in and

around the Guaporé-Mamoré. Strikingly, it does not occur in the western part of the continent with the Andes serving as a border line. This may be further evidence for an Andean vs. Amazonian divide in typological features as postulated by Dixon & Aikhenvald (1999, 8-10). The regional prominence of *FRUSTRATIVE* coincides roughly with the borders of the Tupían expansion. According to Rodrigues (1999, 107), Tupían languages are mostly situated in the Amazon basin. It is very likely that *FRUSTRATIVE* mostly originated in Tupían and diffused with the Tupían expansion. Maps 5.2 and 5.3 illustrate the two clusters of *FRUSTRATIVE* marking in comparison to the almost evenly spread *DESIDERATIVE*.

Turning now to Epistemic Modality, according to a study about the expression of Epistemic Modality as either verbal construction, verbal affix, or else world-wide by Van der Auwera & Ammann (2011), languages in SA do not use verbal constructions to express this category but exclusively either affixes on verbs or “other” means. Their sample includes 25 *SAILS*, 15 of which have exclusively affixation/ cliticization, and 10 have other means which includes particles. In a global perspective, the Americas, Australia, and New Guinea mark Epistemic Modality almost exclusively by non-verbal constructions. Affixation occurs on every continent, but most frequently in the three regions above.

In the present sample, 34 languages mark Epistemic Modality, i.e. either *CERTAINTY* or *DUBITATIVE* or both; by suffixation, cliticization, or particles, or by more than one of those: 14 languages mark Epistemic Modality by suffixation, 12 by particles, 5 by cliticization, four by both particles and suffixes and one by particles and clitics. In the sample, 24 languages mark Epistemic Modality by suffixation or cliticization and 17 by particles (with an overlap of five languages using particles and suffixes/ clitics both). That means that a little over a third of the languages in the sample mark Epistemic Modality by bound forms, a little under a third by non-bound forms (particles) and the rest not at all. Because the present sample does not take verbal constructions into consideration it is not possible to comment upon the claim that there are no languages in SA with Epistemic Modality marked by such. Additionally, the present sample does not investigate whether the languages mark Epistemic Modality exclusively by affixation or whether languages can have more than one means, except for the five languages that have both affixation/ cliticization and particles. These are an exception to Van der Auwera & Ammann’s statement: Hixkaryana, Mosetén, Matses, Tiriyo, and Dâw.

As can be seen in map 5.4, the geographical distribution of Epistemic Marking in SA points to a clear cluster in north-west Brazil and south-east Colombia, including the Vaupés, and several minor clusters in Bolivia and north Peru. The first cluster consists predominantly of unrelated languages but is partly located in the linguistic area of the Vaupés, which is known for its high degree of language change by contact (cf. Aikhenvald 1999). Table 5.9 shows the Epistemic markers of those languages in the sample in and close to the Vaupés area. Interestingly, the *lingua franca* of that region, Nheengatú, features a *DUBITATIVE* particle which is borrowed from Portuguese: *será que* > *seraki* (Da Cruz 2011, 361).

A case of borrowing is the *CERTAINTY* marker in Aymara and Cuzco Quechua: both languages use the suffix *-puni* to indicate that the speaker is relatively certain about the

statement. The fact that a search of neither the Huallaga nor Imbabura Quechua sources revealed a similar form and that Quechua and Aymara have been spoken in the same region for a long time (cf. Adelaar with Muysken 2004) leads to the tentative conclusion that the borrowing took place from Aymara into Cuzco Quechua, although this needs further verification.

(5.86) Aymara (Cerrón-Palomino & Carvajal Carvajal 2009, 201)

sar-xa-puni:-wa

'I will go in any case.' (orig. 'me iré de todas maneras')

(5.87) Cuzco Quechua (Faller 2002, 84)

t'anta-ta-puni irqi-ta-qa qu-rqa-n

bread-ACC-puni child-ACC-TOP give-PST-3

'(S)he certainly gave bread to the child.'

Considering the cluster of Epistemic marking in approximately the Vaupés region, this points toward a likely spread by contact. However, this and the other possible clusters need further research.

Table 5.9: Epistemic marking in/ around the Vaupés

language	Family	CERTAINTY	DUBITATIVE
Cubeo	Tucanoan	-	Cumulative suffixes (FUTURE-DUBITATIVE 3sg. -bU, -bE)
Tariana	Arawakan	-	Clitic (?) = <i>da</i>
Desano	Tucanoan	-	Suffixes - <i>sa</i> , - <i>sia</i>
Nheengatú	Tupían, Guaraní	Particle <i>supi</i>	Particle <i>seraki</i> (from Portuguese)
Puinave	unclassified	Suffix - <i>si</i>	Suffix - <i>tjpa</i> (- <i>ti</i> , - <i>pa</i>)
Dâw	Nadahup	suffix - <i>ih</i> , particle <i>tün</i>	-

I now turn to the geographical distribution of COMMAND types. According to Aikhenvald (2010, 390), the meanings of COMMANDS diffuse relatively easy and thus COMMANDS are prone to be affected by contact, both of forms and patterns. In the present sample, both genealogical and geographical impacts seem to be responsible for the distribution of COMMAND type marking, but no significant amount of COMMANDS due to contact could be established yet.

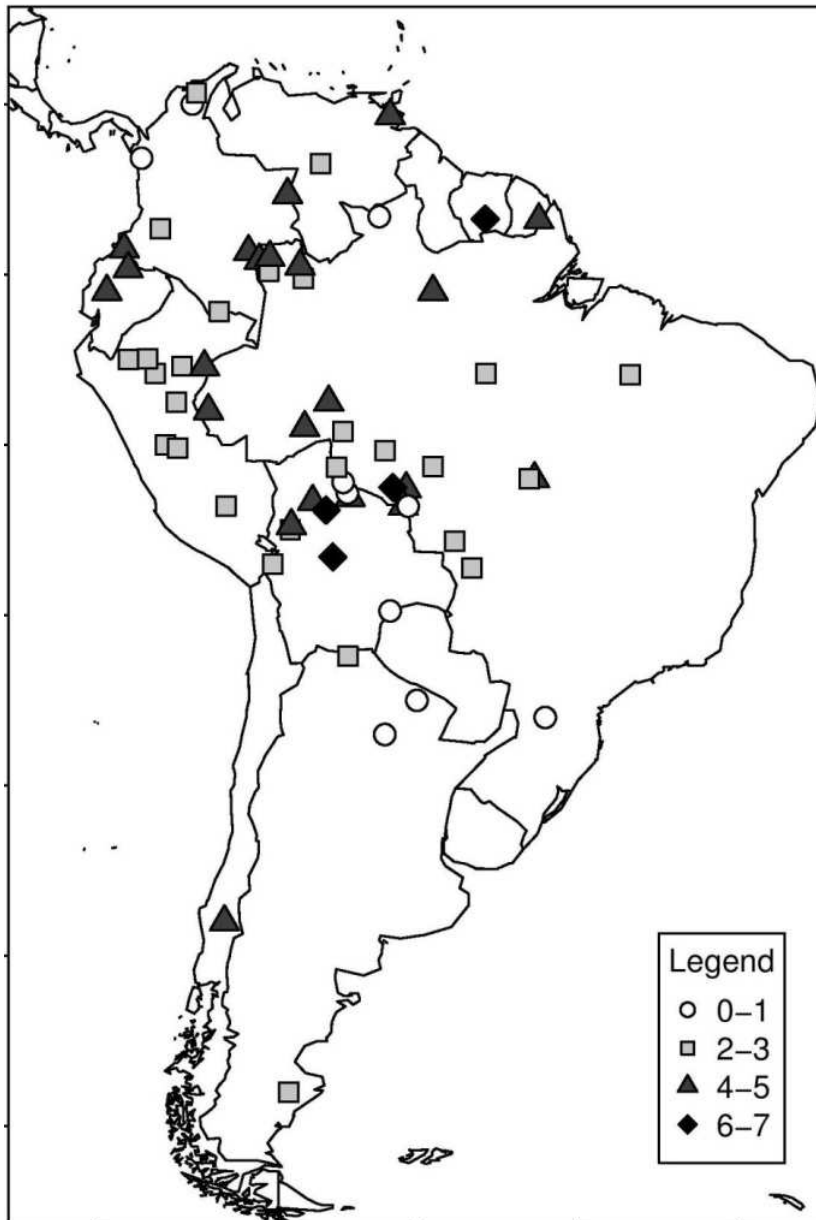
Languages marking at least one command type occur frequently and not restricted to any particular region. As can be seen on map 5.5, languages not marking any COMMAND type are likewise fairly spread over the continent, with a small cluster including both Guaycuruan languages (Mocoví and Pilagá) and two Macro-Gêan languages (Bororo and Kaingang). An areal cluster of HORTATIVE marking includes Rikbaktsa, Kwaza, Mamaindê, and

Sabanê; and of *JUSSIVE* includes Urarina, Aguaruna, and Huallaga Quechua. More research is required to investigate these clusters. *COMMAND* type markers on their own do not point towards clear areal distributions; the following analysis therefore incorporates the results of the *FUTURE Tense* investigation with those of *COMMAND* types.

The geographical distribution of morpho-syntactic Modality marking in the sample is not informative, apart from the clusters mentioned above. The following section investigates whether the genealogical distribution can perhaps contribute more to the distribution of Modal marking in the *SAILS*.

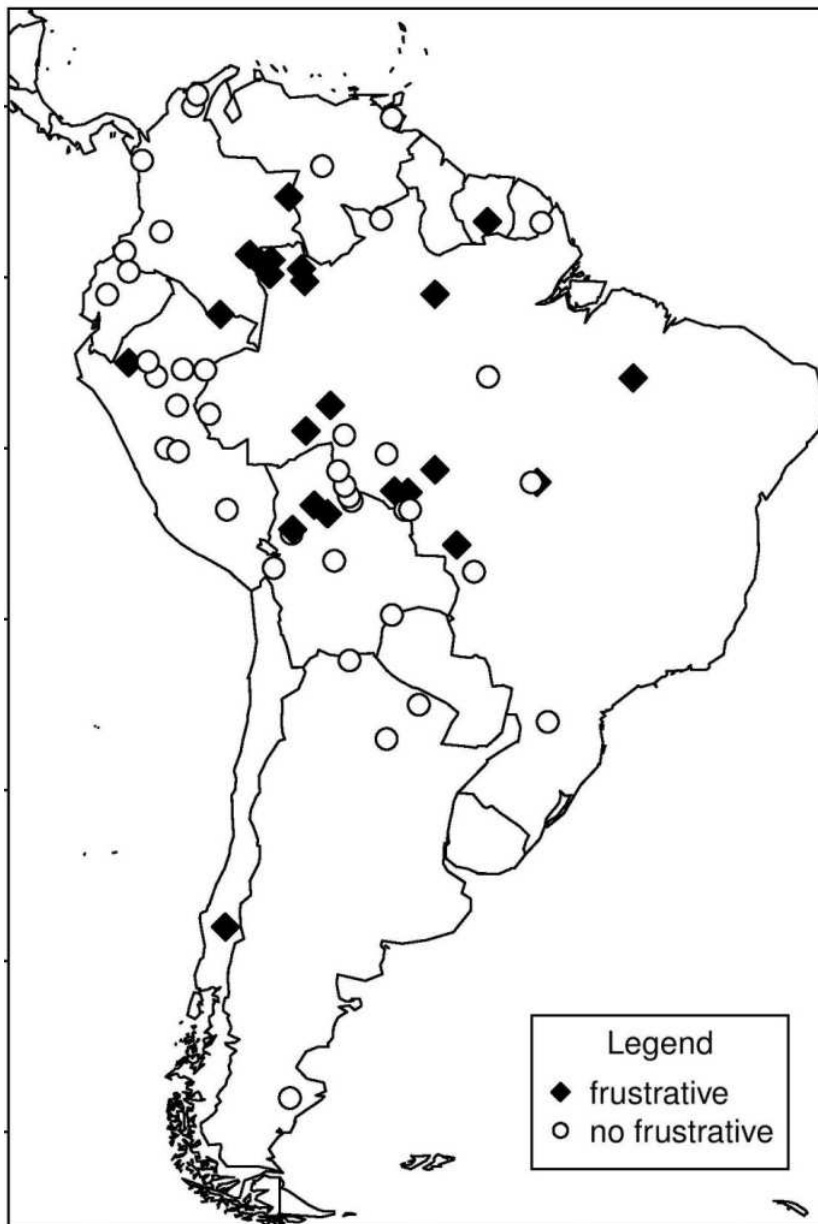
Map 5.1: Distribution of Modals (by number) marked in the sample (excl. COMMAND types)

Number of marked Modals



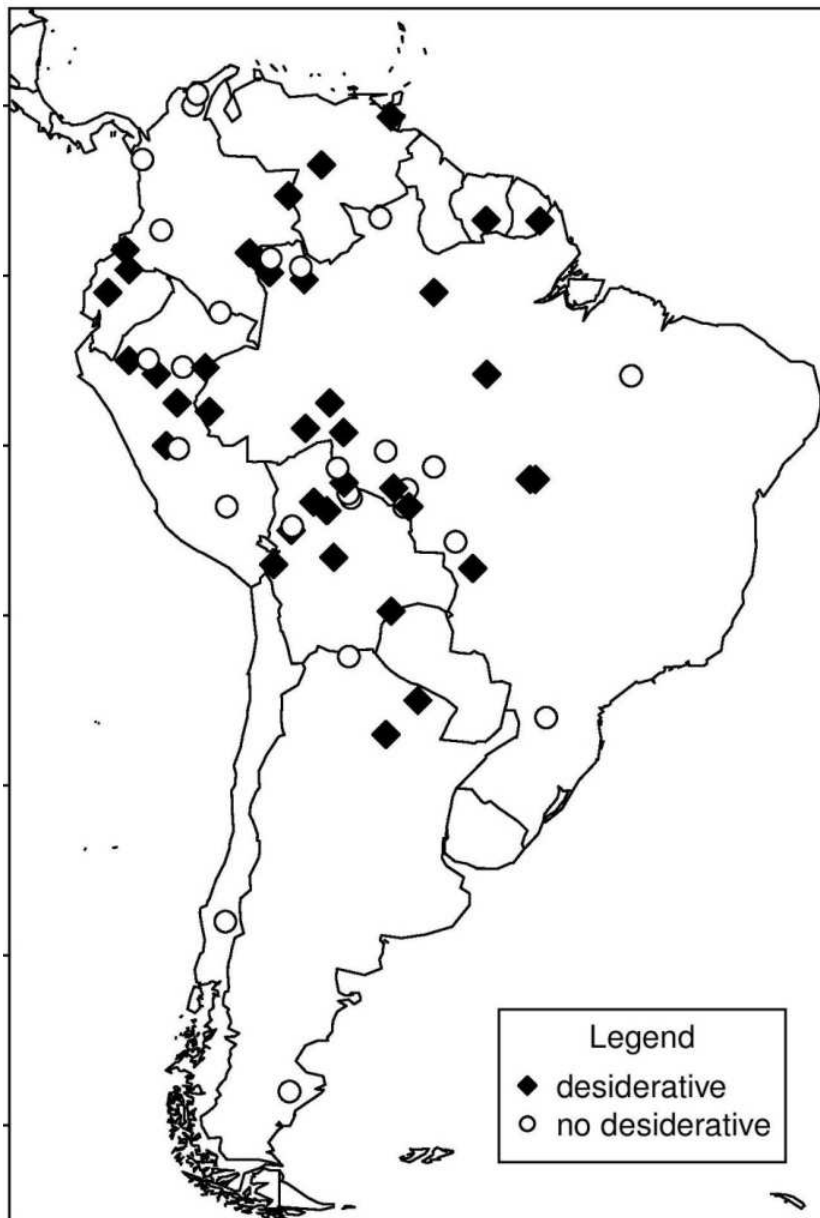
Map 5.2: Distribution of FRUSTRATIVE marking in the sample

Frustrative



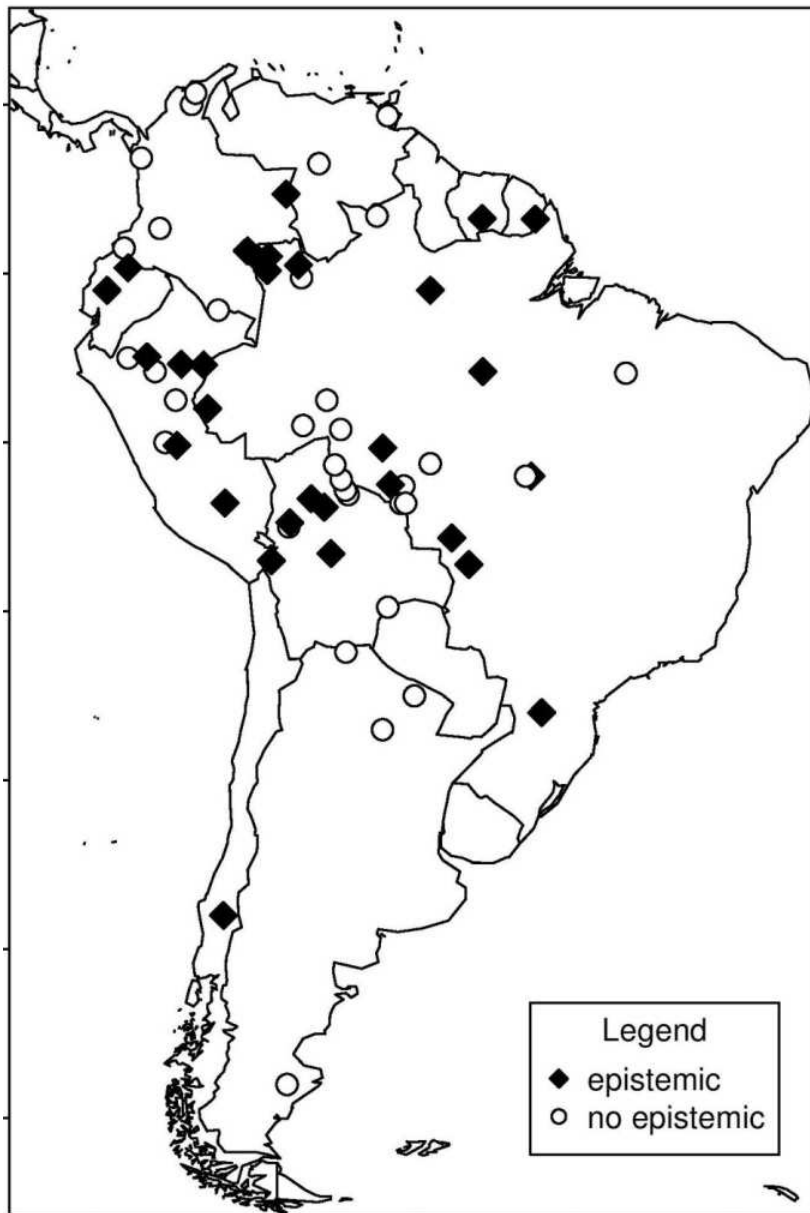
Map 5.3: Distribution of DESIDERATIVE marking in the sample

Desiderative



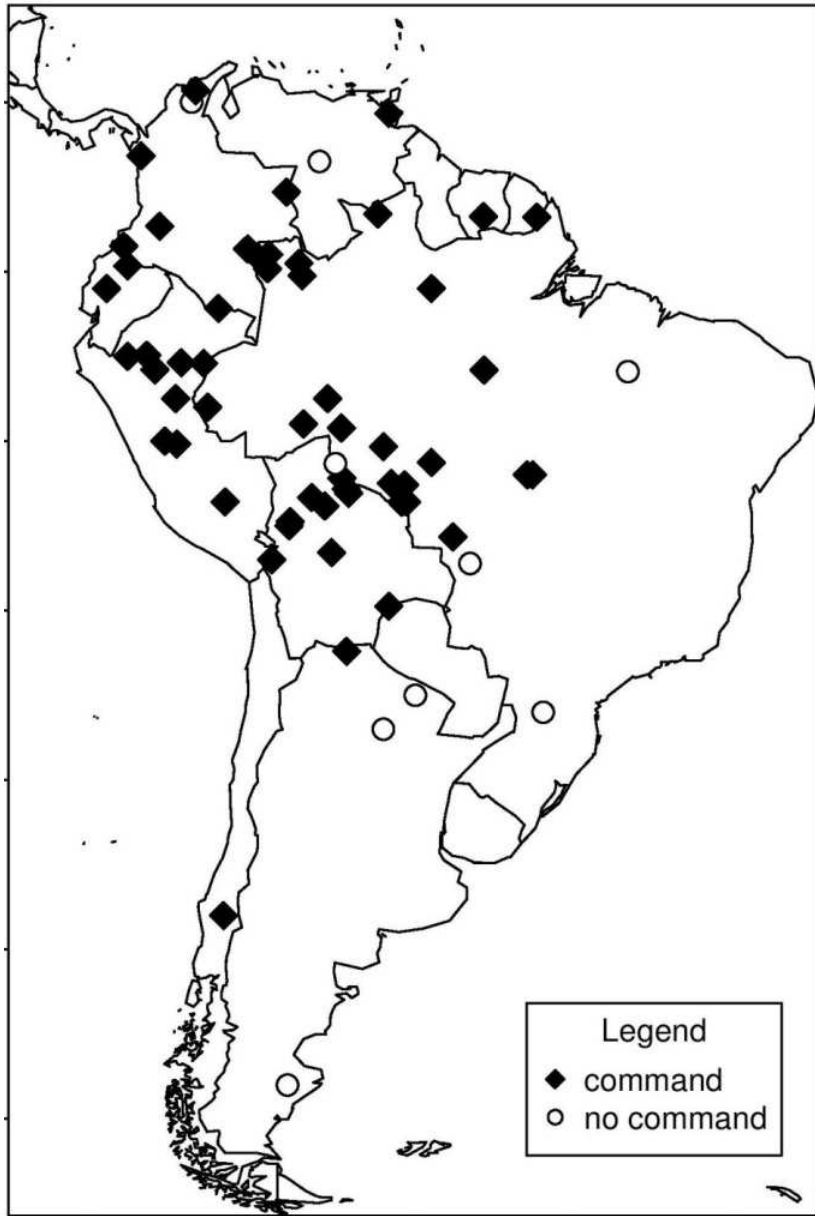
Map 5.4: Distribution of Epistemic Modality marking in the sample

Epistemic Modality



Map 5.5: Distribution of COMMAND marking in the sample

Command marking



5.12 GENEALOGICAL DISTRIBUTION

5.12.1 Introduction

So far the focus has been on distribution of features according to geographic patterns. Now I present the distribution of Modality in the sample with regard to language families. Because the marking of morpho-syntactic Modality is very varied throughout the complete sample, I will concentrate on a few families where more can be said. Other families, for example Macro-Gêan, where no similarity of marking could be found at all, will not be discussed in detail here.

Cognate forms are found in Quechuan, Tupían, Arawakan, and Cariban (see below). In Guaycuruan and Nambikwaran the *DESIDERATIVE* markers are cognates: the suffix *-ake* in Mocoví and Pilagá, and the Sabanê and Mamaindê suffixes *-tan/ -ten* (cf. Mueller 2013). Notable is the absence of *REALIS/ IRREALIS* marking in the Panoan, Nadahup, and Guaycuruan languages. Other language families in the sample, like Arawakan, Cariban, and Tupían, have varied marking. Both Barbacoan languages in the sample have *IRREALIS* marking (but with very different forms). The Guaycuruan and Macro-Gêan languages in the sample stand out as none of them mark any *COMMAND* type.

The following section discusses Modality in Arawakan (5.12.2), Tupían (5.12.3), Cariban (5.12.4), and Quechuan (5.12.5) in more detail.

5.12.2 Arawakan

Because the Arawakan Modality markers in the sample are highly heterogeneous additional data was added (see table 5.11);³³ unfortunately, this did not clarify the picture. According to Aikhenvald (1999, 93), Arawakan languages are characterized by their complex modality systems and that is certainly true in this study as well. For a few cases there is evidence for cognate forms: The *IRREALIS* markers in Baure, Paresi, Apurinã, and Terêna may be reflexes of the proto-form **-sia* for *FUTURE/ IRREALIS/ POTENTIAL* (Payne 1991, 381). The Epistemic Modality markers in Paresi, Dení, and Waurá certainly look related, as do the *FRUSTRATIVE* markers in Matchiguenga, Nomatsiguenga, and Ashaninca/ Asheninca. *FRUSTRATIVE*, listed by Aikhenvald (1999, 94) as one feature of those Arawakan languages that have rich Modal systems, occurs in Paresi, Apurinã, Terêna, Amuesha, Ashaninca/ Asheninca, Nomatsiguenga, Matchiguenga, and Tariana (see also section 5.7.2).

Wise (1991, 614) states that usually *IMPERATIVE* in Arawakan does not differ from second person declarative forms. Aikhenvald (1999, 94) adds that in general there is only one *IMPERATIVE* per language. While in the languages in the sample *COMMAND* marking is in general not obligatory, there are certain forms that do mark *IMPERATIVE*. Tariana employs almost the whole range of *COMMAND* types except for *JUSSIVE*, and all the other Arawakan languages mark at least one of them but not more than three. It is even possible to identify two cases of

³³ Additional data for Terêna, Wapishana, Bhawana/ Chiriana, Achagua, and Waurá from Aikhenvald (1999), for Dení, Waurá, and Terêna from Derbyshire (1991), and for Amuesha, Nomatsiguenga and Ashaninca/ Asheninca from Wise (1991).

probable genetic inheritance: the JUSSIVE particle *ta* in Baure and the HORTATIVE enclitic *=da* in Tariana, and the HORTATIVE particle *ši* in Baure and the proximate suffix *-si* in Tariana. Interestingly Desano, which is Tucanoan but in contact with Tariana, also has a COMMAND suffix: HORTATIVE (permissive) *-si*. Aikhenvald (ibid.) suggests that the large COMMAND type system in Tariana is due to areal diffusion from Tucanoan into Arawakan, but the fact that the marker *-si* also occurs in another Arawakan language points toward borrowing from Arawakan into Tucanoan rather than the other way round. Additionally, the marker may have developed from the FUTURE marker *-si* (Aikhenvald 2010, 381), as noted. On the other hand, the Tariana marker *-si* occurs exclusively with second person whereas the Desano marker refers to first person which suggests that the similarity of the forms may be due to chance rather than contact. The Tariana IMPERATIVE suffix *-ya* is definitely borrowed from Tucanoan, though (Aikhenvald 2003, 379; Aikhenvald 2010, 381). The exact same form occurs in COMMANDS of Barasano and Tuyuca *-ya* (Aikhenvald 2012, 187-189) as well as Desano (Miller 1999, 72). The FUTURE suffixes *-ya* in Embera (Chocoan) and Puinave (unclassified) are possibly related as well. In that case the form would have spread from Tucanoan into Tariana (Arawakan), Embera (Chocoan), and an unclassified language, but that has yet to be proven.

Table 5.10: COMMAND type marking in Arawakan

LANGUAGE	JUS	POL COM	COL COM	HORT	PROH	IMP
Yanesha'	-	-	-	-	<i>-ats</i>	-
Tariana	-	<i>=thara</i>	Detrimental clitic <i>=tupe</i> (2+3)	<i>=da</i>	<i>mhaida</i>	Proximate suffix <i>-si</i> , distal suffix <i>-kada</i> , postponed <i>-wa</i> , secondhand <i>=pida</i> , <i>-ya</i> (borrowed from Tucano)
Paresi	-	<i>maika</i>	-	-	<i>awa</i>	-
Apurinã	-	-	-	<i>(h)amo</i>	-	<i>-poka</i> (rare, used to be polite only)
Baure	<i>ta</i>	-	-	<i>ši</i>	-	<i>-no</i>

5.12.3 Tupían

The Tupían languages mostly vary as to which Modal category they mark, but there are a few sets of cognates, most prominently the FRUSTRATIVE. The FRUSTRATIVE markers and forms with related meanings in the sample and as found in Jensen (1998, 538-539) seem to be related, probably with two different sources. The Nheengatú and Kamaiurá markers *yepe* and *jepe*, respectively, occur as Epistemic Modals in other Guaraní languages (Dietrich 2006, 78), which points towards a grammaticalization from Epistemic Modality to FRUSTRATIVE. Reflexes of the IMPERATIVE proto-forms in Guaraní branch **e-* (sg), **pe-* (pl.) (Jensen 1998, 525) are found in Tapiete, Tupinamba, and Emérrillon, but Cocama-Cocamilla has a completely different IMPERATIVE form. The Tupían language Karo which is not of the Guaraní branch has a probably related COLLECTIVE COMMAND prefix *pe?*-. Likewise, the Awetí (non-Guaraní) prefixes *i-* (sg) and *pej-* (pl) (PERMISSIVE) (cf. Drude 2008, 80) may be cognate. The Tupían (Guaraní) proto-PROHIBITIVE **eme* (Jensen 1998, 549) is found in Kamaiurá and possibly Emérrillon.

For a discussion of the DESIDERATIVE cognates in Tupían see Mueller (2013).

Table 5.12: FRUSTRATIVE in Tupían³⁴

	FRUSTRATIVE 1	FRUSTRATIVE 2
Mekens	<i>etaop</i> (?)	-
Kamaiurá	<i>jepe</i>	-
Nheengatú	<i>yepe</i>	-
Awetí	<i>tepe</i> ³⁵	-
Wayampi	<i>tite</i>	<i>mijã</i> 'previously, but not any more'
Urubú-Kaapor	<i>tipe</i>	-
Tupinambá	-	<i>-biã</i>
Guarayu	<i>tēi</i>	<i>vjã</i> 'unreal'
Guajajára	<i>tezyz</i> 'frustration', <i>zepe</i> 'incomplete success'	<i>miamo</i> 'in vain'

³⁴ Wayampi, Urubú-Kaapor, Tupinambá, Guarayu, and Guajajára data from Jensen (1998).

³⁵ In Drude (2008, 93).

Table 5.13: Modality in Tupían

	Sateré- Mawé	Karo	Karitiana	Mekens	Tapiete	Cocama- Cocamilla	Emérillon	Kamaiurá	Nheengatú
REA	-	-	-	-	-	-	-	-	-
IRR	-	-	i-	pegat	-	-	itjẽ-ite	ram	-
CERT	-	-	-	ebõ, êp, te	-	=tin	-	-	supi
DUB	sio, som	pə, iʔkiy	-	kẽra, toët	-	=taka, =ray	-enã ³⁶	inip ~ nip	seraki
DES	-teran	-	-wak	kot ke, kot kaat	-se, -(i)sha	-	-tanẽ	-potat, -wej	-
INT	-	-	-	-	-	-	-	-	-
FRUST	-	-	-	etaop	-	-	-	jepe	yepe
PRP	-	nāt	-	-	-	=ra	t-	-	arã
POT	aru	-	-	eteet	-	-	-	in	-

Table 5.14: COMMAND type marking in Tupían

	IMP	HORT	JUSS	POL COM	COL COM	PROH
Sateré- Mawé	o, to	-	-	-	-	mei'o/ tei'o
Karo	-	-	-	-	peʔ- (1/3)	yahmãm
Karitiana	-a	-	-	-	-	-
Mekens	-	-	-	-	-a (2sg, pl/ 1pl)	-bõ
Proto-form	*e- (sg), *pe- (pl)	-	-	-	-	*eme
Tapiete	e-	-	-	-	t-/ti- (1/3)	awi
Cocama- Cocamilla	yawa	yapay	-	-puri, tsani/ tsaniuri	-	ina
Emérillon	e- (sg), pe- (pl)	-	-	-	t-	mame-mamen
Kamaiurá	-	-	-	-	ta=	-un(e), panen
Nheengatú	-	-	-	-	-	=te
Tupinamba	e- ³⁷	?	?	?	?	?

5.12.4 Cariban

In Cariban we find evidence for two cases of cognates: The DESIDERATIVE markers in Hixkaryana, Tiriyo, and Panare, and the PURPOSIVE in Hixkaryana, Tiriyo, Wai Wai, and Makushi (data for the latter two are from Derbyshire 1999). Gildea (1998, 138-140) reconstructs a proto-Cariban purpose construction involving the attributive *me/pe and nominalizer *-topo which occurs in Tiriyo, Hixkaryana, and Panare, but the former two together with

³⁶ Similar Epistemic Modality markers in Suruí -*iná* and *ená* (Rodrigues 1999, 119).

³⁷ In Jensen (1999, 148).

Wai Wai and Makushi also exhibit special PURPOSIVE markers that likewise are cognates. Gildea (1998, 141-151) demonstrates that the participial **-ce* is very frequent in Cariban and oftentimes described as “purpose-of-motion”, which fits the description of at least the Hixkaryana and Tiriyo markers (see section 5.6 for examples). It is very likely that the DESIDERATIVE markers in Hixkaryana and Tiriyo are also cognate with **ce*.

As Derbyshire (199, 37) notes: “[t]here is another set of suffixes for the imperative [differing from the indicative] (including hortatory mood); these have components of number and motion, and differ in form according to the person of the subject”. Because data for Panare is scarce this can only be confirmed for Tiriyo and Hixkaryana.

In general, Modality marking in Cariban looks homogeneous only to a small degree. INTENTIONAL is marked only by Tiriyo, and the markers of IRREALIS, CERTAINTY, and FRUSTRATIVE do not seem to be genealogically related. Tiriyo is the language with the most marked Modality categories in the sample, and Hixkaryana likewise ranks in the top third. This is explained by the high degree of fusion in the Cariban languages in the sample; Epistemic Modality, FUTURE, and PERFECTIVE/ IMPERFECTIVE are expressed by cumulative morphemes in Tiriyo (cf. Meira 1999) and in Hixkaryana Epistemic Modality and NONPAST forms occur in cumulative morphemes. Because the sources for Panare are not comprehensive on TAME marking, it is expected that, after more data become available, there will be a similar high degree of Modality marking and fusion.

Table 5.15: Modality marking in Cariban

	Hixkaryana	Tiriyo	Panare	Wai Wai	Makushi	Proto-form
REA	-	-	-	?	?	?
IRR	-	<i>mo</i>	<i>t-</i>	?	?	?
CERT	<i>mpini</i>	<i>-e</i>	-	?	?	?
DUB	<i>-yano/ -yatxowi; na, mpa</i>	<i>-ne/ -nē, tahke(ne), tahkarē, tahkara</i>	-	?	?	?
DES	<i>-txahke, fe</i>	<i>se/ fe</i>	<i>-jte</i>	?	?	<i>*-ce (?)</i>
INT	-	<i>-ta</i>	-	?	?	?
FRUST	<i>haryhe</i>	<i>re(pe)</i>	-	?	?	?
PRP	<i>hori, -so/ -xe</i>	<i>-se</i>	-	<i>-so/ -fi</i>	<i>-i/ -se</i>	<i>*-ce</i>
POT	-	-	-	?	?	?

Table 5.16: COMMAND type marking in Cariban

	Hixkaryana	Tiriyo	Panare
IMP	<i>-ko/ -txoko</i>	<i>-(kē), -ta, mii/ miiiko(mo)</i>	-
HORT	<i>-si/ -xe, -sinye/ -xenyē</i>	<i>-ne (pl)</i>	-
JUSS	<i>-(haka?)</i>	-	-
POL COM	-	-	-
COL COM	-	<i>ēwē(h) (3/1 sg)</i>	-
PROH	-	-	-

5.12.5 Quechuan

Modality marking in Quechuan is highly homogeneous. Almost all of the Modality markers in the three Quechuan languages in the sample and an additional set of four (Ayacucho, Tarma, Salasaca, and Pacaraos Quechua)³⁸ can be brought into relations: The IRREALIS markers in Huallaga, Imbabura, and Cuzco Quechua and the POTENTIAL marker in Ayacucho and Tarma Quechua are identical, and the DUBITATIVE markers in Imbabura, Cuzco, Salasaca, and Pacaraos Quechua are certainly related (possibly also the IRREALIS in Ayacucho). The DESIDERATIVE markers as found in Imbabura and Cuzco Quechua grammaticalized from the verb *muna* ‘want’ (cf. Mueller 2013).³⁹ The Epistemic marker *-puni* in Cuzco Quechua has an Aymaran homonym (Adelaar with Muysken 2004, 291).

The same homogeneity holds true for the set of COMMAND type markers. The IMPERATIVE suffix *-y* as well as the JUSSIVE suffix *-chun* occur in almost all languages in table 5.18.

Table 5.17: Modality marking in Quechuan

	Huallaga Quechua	Imbabura Quechua	Cuzco Quechua	Ayacucho Quechua	Tarma Quechua
REA	-	-	-	?	?
IRR	<i>-man</i>	<i>-man</i>	<i>-man</i>	<i>-ču</i>	?
CERT	-	-	<i>-puni, -mi</i>	?	?
DUB	-	<i>-chá(ri)</i>	<i>-chus, -chá⁴⁰</i>	?	?
DES	<i>-na:</i>	<i>-naya</i>	-	?	?
INT	-	-	-	?	?
FRUST	-	-	-	?	?
PRP	-	<i>-chun/ -ngapaj</i>	-	?	?
POT	-	-	-	<i>-man</i>	<i>-man</i>

Table 5.18: COMMAND type marking in Quechuan

	Huallaga Quechua	Imbabura Quechua	Cuzco Quechua	Ayacucho Quechua
IMP	Suffix <i>-y</i> , (<i>-nki</i> 2pl FUT)	<i>-y</i> (sg), <i>-chi</i> (pl)	<i>-y</i>	<i>-y</i>
HORT	<i>-shun</i>	<i>-shunchi</i>	-	<i>-?</i>
JUS	<i>-chun</i>	<i>-chun</i>	-	<i>-čun/ -sun</i>
COL COM	-	-	-	?
POL COM	-	-	-	?
PROH	-	-	-	?

³⁸ Additional data from Adelaar with Muysken (2004).

³⁹ Incidentally, the verb *muna* ‘want’ also occurs in Aymara (cf. Adelaar with Muysken 2004, 272, ex. 130), but no DESIDERATIVE that grammaticalized from it.

⁴⁰ See section 6.3.3 Quechuan for Evidential/ Epistemic *-chá*.

5.13 STABILITY OF MODALITY

This section evaluates the stability of Modal features and draws upon the studies by Wichman & Holman (2009) and Wichmann & Kamholz (2008) (see chapter 3.10 for an introduction to stability of typological features and the stability of Tense, 4.11 for Aspect, 6.6 for Evidentiality, and 7.5 for a final discussion). The Modal features in Wichmann & Holman (2009) include “Situational Possibility”, “Epistemic Possibility”, “Overlap between Situational and Epistemic Modal Marking”, “The Morphological Imperative”, “Prohibitive”, and “imperative-Hortative Systems”, all of them ranked relatively unstable or very unstable, and “Purpose Clauses” which are ranked relatively stable. In general, the present data confirm that Modal categories are not stable and prone to language internal change as well as to diffusion. It has already been suggested that the (in)stabilities of Tense and Aspect determine each other, and the same is possibly true for Modality as well, because of the demonstrated relationship between at least FUTURE and many Modal categories (e.g. COMMANDS, INTENTIONAL, IRREALIS, etc.). Additionally, Modal features are characterized by their susceptibility to pragmatic factors, such as politeness in COMMANDS, and these features tend to be unstable in general (Wichmann & Holman 2009, 33).

5.14 SUMMARY

This chapter investigated the typological, geographical, and genealogical patterns of Modality marking in the sample. The languages were analyzed according to REALIS/IRREALIS, INTENTIONAL, CERTAINTY, DUBITATIVE, POTENTIAL, FRUSTRATIVE, DESIDERATIVE, PURPOSIVE, and the following COMMAND types: IMPERATIVE, HORTATIVE, JUSSIVE, COLLECTIVE COMMAND, POLITE COMMAND, and PROHIBITIVE.

Results show that IMPERATIVE is the most frequently marked category (45 languages), followed by DESIDERATIVE (39), PROHIBITIVE (30), DUBITATIVE (28) and IRREALIS (28), PURPOSIVE (24), FRUSTRATIVE (23) and HORTATIVE (23), CERTAINTY (20), POTENTIAL (17), COLLECTIVE COMMAND (16), INTENTIONAL (13), POLITE COMMAND (11) and JUSSIVE (11), and REALIS (3). The overall degree of suffixation is very high. Only one language has no Modality marking (Chimila), but none marks the maximum possible number (15 including COMMAND types; nine excluding COMMAND types). Disregarding COMMAND types, Tiriyo marks the most Modal categories (7).

Several regional clusters of Modal marking occur in the sample. The Guaporé-Mamoré area features a considerable variation of high genealogical diversity; on the contrary, the Vaupés region is relatively homogeneous in terms of which Modal category is marked, although the markers themselves do not show any similarities. FRUSTRATIVE has been shown to occur only east of the Andes in what roughly constitutes Amazonia, and throughout all language families, but without genealogical signal except for Tupían. Tupían prominently features FRUSTRATIVES which may have two different origins, and possibly contact with Tupían led to the spread of FRUSTRATIVES in Amazonia. Whereas a single language family was probably responsible for the spread of FRUSTRATIVE, the DESIDERATIVE was developed by several language families, which explains its wider geographical spread than the FRUSTRATIVE one.

Epistemic Modality occurs particularly frequently in north-west Brazil and south-east Colombia, including the Vaupés, and several minor clusters in Bolivia and north Peru. COMMANDS occur frequently in the sample in all regions and families, with the exception of Guaycuruan and Macro-Gêan. Many COMMAND type markers apparently grammaticalized from the same origins as FUTURE markers, or from FUTURE markers.

Modality marking varies greatly according to family. Quechuan is very homogeneous in Modality and COMMAND type marking, with cognates for almost all categories. On the other hand, Macro-Gêan is very heterogeneous and unified only by its lack of COMMAND type marking. Guaycuruan has no COMMAND type marking as well. The other language families rank in between these two extremes, being relatively heterogeneous with a few sets of cognates. Tupían has FRUSTRATIVE cognates; DESIDERATIVE occurs in cognates in virtually all families; COMMAND type cognates in Cariban and Tupían can be connected to established proto-forms.

Modal features are unstable, both in the present sample and Wichmann & Holman's (2009) study, because on their relationship with Tense, which is also unstable, and its susceptibility to pragmatic factors.

6. EVIDENTIALITY

6.1 INTRODUCTION

This chapter investigates the distribution of the grammaticalized marking of information source, also termed Evidentiality. The notion of Evidentiality is not new to typological research and has been the topic of various studies on the distribution of the phenomenon, its function within grammatical systems, and its various semantic, pragmatic, and discourse characteristics. For overviews of the history of Evidentiality in linguistics see e.g. Willett (1988), Dendale & Tasmowski (2001), and Aikhenvald (2004). Despite this world-wide interest in marking of information source, Evidentiality is far from fully explored. For example, it continues to be the topic of hot debates about whether or not Evidentiality is part of Modality or indeed exists as a category in its own right.

The present study focuses on the morpho-syntactic marking of Evidentiality in SAILS, based on carefully defined semantic components. It does not attempt to resolve all the points of debate, but instead offers an empirical overview in the SAILS sample, the results of which will be analyzed on the basis of recent literature. The categories chosen are those that recur in recent studies as highly relevant, in terms of frequency as well as meaning. Before I present the individual definitions, however, a general note about Evidentiality is in order.

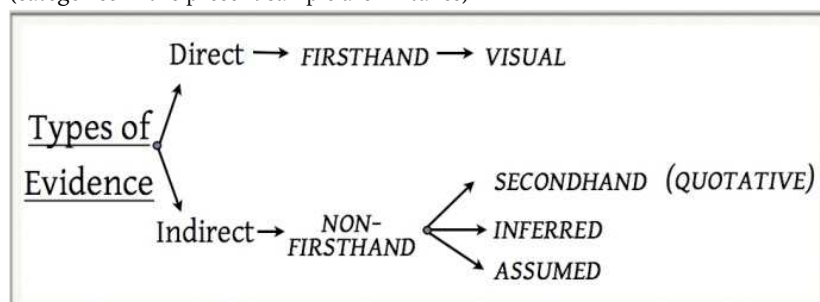
The concept of marking information source (e.g. Evidentiality) has been defined by Aikhenvald (2004, 3) as follows:

“Evidentiality is a linguistic category whose primary meaning is source of information (...) this covers the way in which the information was acquired, without necessarily relating to the degree of speakers’ certainty concerning the statement or whether it is true or not (...). To be considered as an evidential, a morpheme has to have ‘source of information’ as its core meaning; that is, the unmarked, or default interpretation”.

The questionnaire follows this approach and the definitions outlined in section 6.2 are also closely modeled on the ones by Aikhenvald.

The dominant function of any Evidential is that it signals how the information referred to by the speaker was acquired. Following this, Evidentiality can pertain to a number of acquisition routes, starting at a rough direct vs. indirect witness opposition and branching into refined meanings such as witness by vision or hearing, or inference based on indirect evidence. The logical internal structure of Evidentiality is basically the same throughout Evidential studies and will also be used as model here. The following figure presents a simplified version of the semantic internal structures of Evidentiality based on Willett (1988) and Aikhenvald (2004). Evidential categories included in the questionnaire are in italics.

Figure 6.1: Types of Evidentiality in this study based on Willett (1988) and Aikhenvald (2004) (categories in the present sample are in italics)



This thesis does not investigate all types of Evidentiality presented by Willett, but adds two extra distinctions (*QUOTATIVE* and *ASSUMED*).

Whereas the internal organization of Evidentiality is generally agreed upon by typologists, they are less unanimous about grouping the whole category in relation to other grammatical categories. There are two opposing sides: those that view Evidentiality as a subgroup of Modality, and those that see it as a category outside of, but connected to, Modality (and Tense and Aspect). The first opinion is also the traditional one, endorsed by e.g. Willett (1988) and Palmer (2001), whereas the second one is much more recent and advocated by e.g. Aikhenvald (2004) and de Haan (2005). This debate is intertwined with the one about the exact relationship between Evidentiality and Epistemic Modality, i.e. are these conjunct (the same), disjunct (apart), or overlapping systems (cf. Dendale & Tasmowski 2001, 341f.). As in this study they are presented as separate categories, I will now discuss this point in more detail.

Palmer (2001, 8) puts Evidentiality on a par with Epistemic Modality, the two categories forming a binary distinction within the category of Propositional Modality. His main argument is that a speaker makes a judgement about the factual status of the proposition (Epistemic) when indicating the acquisition source (Evidential), so these notions are inextricably linked: e.g. inferring from an indirect source always results in a relatively uncertain statement. Opponents argue that marking the source of information does not necessarily involve judgement and that it indeed can be related with, but is independent of, Epistemic Modality. Just as judging the factuality of a proposition does not have to involve source of information, giving a source can be neutral as to judgment.

De Haan (2001, 208) argues that a contingent link between Epistemic Modality and Evidentiality does not exist, but rather is speaker-dependent. Aikhenvald (2004, 7-8, footnote 1) argues that “the occasionally existing link between some evidential choices and the expression of certainty or uncertainty [...] is then mistaken for a universal”. She demonstrates that languages indeed can have distinct (disjunct) systems of Evidentiality and Epistemic Modality, that these exist independently, and can even co-occur in the same

clause (ibid. 258-260). The sole existence of languages that have disjunct systems supports the view that Evidentiality \neq Modality.¹

While the present study acknowledges that there can be relations between Evidentiality and Epistemic Modality, they are not seen as dependent on each other, or as part of the same category. Whereas Epistemic Modality is concerned primarily with the speaker's judgment of the proposition, Evidentiality refers to a concept completely independent from the speaker's judgment, or, in De Haan's terms (2005, 380; author's emphasis): "Evidentiality asserts the evidence, while epistemic modality *evaluates* the evidence".

The following section (6.2) outlines the Evidential categories chosen for this study and their definitions. A primary reason why they were chosen is based on the results from Aikhenvald (2004, 63), who states that six recurrent semantic parameters can be found throughout the languages of the world: visual, non-visual, inference, assumption, hearsay, and quotative.² These parameters were chosen here as well with the difference that non-visual was substituted for NON-FIRSTHAND to allow for a broad FIRSTHAND/ NON-FIRSTHAND distinction. Section 6.3 presents my results. General findings are discussed in 6.3.1. Section 6.3.2 discusses the Evidential systems as occurring in the sample, and 6.3.3 the distribution of Evidential markers in relation to genealogical and geographical patterns. The results are summarized in section 6.4.

6.2 DEFINITIONS

6.2.1 FIRSTHAND and VISUAL Evidentiality

FIRSTHAND information refers to the speaker being the direct witness to the marked event; this can be through vision or other sensory experiences like hearing, touch, smell, and taste. A marker in this category can refer to any of these. A marker that exclusively refers to visual witness is not counted as FIRSTHAND, but as VISUAL instead. See an example from Karo for a VISUAL particle (*topə*) in (6.1), and from Hup for a FIRSTHAND marker that excludes VISUAL in (6.2a) and (6.2b). In Hup, the absence of any Evidential usually expresses VISUAL (Epps 2008, 643).

(6.1) Karo (Tupían; Gabas Jr. 1999, 266)

<i>péj</i>	<i>ʔe-t</i>	<i>topə</i>	<i>to=at</i>	<i>macaʔit</i>	<i>wĩ-a</i>
white.man	AUX-IND1	be.seen	3R=POSS	pet	kill-GER

'(It was seen that) the white killed his pet.' (used in a situation where the speaker went to the white man's house and saw him killing his pet)

¹ Cf. Da Cruz (2011, 353) who claims that the REPORTATIVE in Nheengatú has no inherent Epistemic value.

² Based on a convenience sample of 500 languages chosen for availability of data (A. Y. Aikhenvald 2004, xii).

- (6.2) Hup (Nadahup; Epps 2008, 646, 643)
- (a) *náciya pæ-cíw-íy=hō*
 boat go.upriver-COMPL-DYN=NVIS
 ‘The boat already went upriver.’ (the speaker heard but did not see it)
- (b) *g’əh náw=hō*
 sweet good=NVIS
 ‘It’s nice and sweet!’ (tasting something)
- (c) *ʔok-níh key-ham-g’et-yíʔ-ay=∅ tih=ʔáŷ-áh!*
 move-NEG see-go-stand-TEL-INCH=VIS 3SG=F-DECL
 ‘She was just standing there looking, without moving!’ (speaker witnessed event)

6.2.2 NON-FIRSTHAND Evidentiality

A NON-FIRSTHAND marker refers to the opposite of direct witness. As this is in direct opposition to FIRSTHAND, it is expected that this includes everything not experienced through the senses, although often this means everything not acquired by visual information, and therefore could also be called non-visual. This meaning comprises all the notions in figure 6.1 that depend on the node ‘indirect’, i.e. SECONDHAND, INFERRED, ASSUMED, and QUOTATIVE. However, a marker that exclusively marks one of these is not counted as NON-FIRSTHAND, but as the respective submeaning. Only when a marker generally encodes indirect witness, having two or more functions of SECONDHAND, INFERRED, ASSUMED, or QUOTATIVE, it is a NON-FIRSTHAND marker (this is similar to IMPERFECTIVE, which is defined as marking at least two of the meanings of HABITUAL, CONTINUATIVE, ITERATIVE). NON-FIRSTHAND markers are not the same as SECONDHAND markers, but SECONDHAND is a possible meaning of NON-FIRSTHAND. For example in Tapiete, the NON-FIRSTHAND marker *-nda* refers to verbal report, hearsay, but also inference:

- (6.3) Tapiete (Tupían, Guaraní; González 2005, 252, 253, 254)
- (a) *ou-nda arka’e ko-pe*
 3:come-NFRST long time ago DEM-LOC
 ‘(they say that) (s)he came here a long time ago.’
- (b) *yáwa-da ha’e-ipi iko*
 tiger-NFRST (s)he-LOCMOV be
 ‘(It is said that) the tiger lives over there.’
- (c) *ñaimba-nda ndu-su-e*
 dog-NFRST 2SG-bite-REC
 ‘It seems it is the dog (the one that) bit you.’

6.2.3 SECONDHAND and QUOTATIVE Evidentiality

SECONDHAND refers to information that is acquired through verbal report from someone else, also called ‘reportative’. This includes third-hand, fourth-hand, etc. and also the meaning often identified as hearsay or rumor. SECONDHAND information can, but does not have to,

include the meaning of QUOTATIVE. QUOTATIVE is a specialized function of SECONDHAND. A QUOTATIVE occurs in a direct speech construction with a source that is usually directly identified or identifiable by context. Semantically, the source of SECONDHAND information is not important, important is that the information was acquired through somebody else. With a QUOTATIVE the source is important, as it indicates not only that the information was acquired through somebody else, but also by whom exactly. SECONDHAND markers are often translated into English with ‘it is said...’ or ‘they say...’, and are typically used for narratives or hearsay where the speaker repeats a story or information that he has heard from another person, but may not remember where or who from. (6.4) provides an example of the SECONDHAND marker *pa* in Cavineña):

- (6.4) Cavineña (Tacanan; Guillaume 2008, 644):
amena [*tuna_s* *nawi-karama* *ju-ya=ju*] =*pa=tuna_A* *katsa-kware*.
 BM 3PL bathe-DESID.NEG be-IMPV=DS=REP=3PL (-ERG) beat-REM.PAST
 ‘It is said that when they (our ancestor’s children) refused to bathe, they (our ancestors) would beat them.’

SECONDHAND and QUOTATIVE often overlap. In Cavineña, the SECONDHAND marker *pa* can also be used with an identified source, as ‘the foreign lady’ in (6.5):

- (6.5) Cavineña (Tacanan; Guillaume 2008, 645)
runeshi =*pa=ekwana_o* *gringa=ra_A* *iya-mere-ya* *avion=eke*.
 monday=REP=1PL foreign.lady=ERG put-CAUS-IPFV plane=PERL
 ‘The lady said that she will have someone (pilot) taking us (back to our community) by plane on Monday.’

Pa is nevertheless coded as SECONDHAND, because it occurs with indirect speech and a majority of the examples do not seem to identify the source. Furthermore, Cavineña has a separate construction for quotation that has no QUOTATIVE marker but involves repeating of direct speech followed by the deictic adverb *jadya*:

- (6.6) Cavineña (Tacanan; Guillaume 2008, 777)
amenat uekedyá [*“Pa-kwa=dyai-ke_s!”* *jadya*] *ju-kware*.
 BM then HORT.SG-GO=FOC 1SG-FM thus be-REM.PAST
 ‘But then (as nobody wanted to go) I said: “I’ll go!”’

As QUOTATIVE markers are just a subcategory of SECONDHAND, they are expected to be rarer than SECONDHAND markers in the sample.

6.2.4 INFERRED and ASSUMED Evidentiality

Both INFERRED and ASSUMED Evidential markers encode information that has been obtained through indirect measures, i.e. belong to the NON-FIRSTHAND category, just like SECONDHAND and QUOTATIVE. With an INFERRED marker, a speaker infers information on the basis of sensory input. This is easily mistaken for FIRSTHAND information, which is also based on sensory input, but with the difference that with an INFERRED marker the event itself is not directly witnessed. In (6.7), the speaker infers that it rained because the ground is wet, i.e. he did not see it rain but it he sees the direct result. In (6.8) the speaker deduces from seeing the dead prey that the other person must have killed it, although he did not witness the killing:

- (6.7) Kwaza (unclassified; Van der Voort, 421)

a'we-hỹ-tehere

rain-NOM-INFR

'it must have rained' (I did not see it rain, but the ground is wet)

- (6.8) Karo (Tupían; Gabas Jr. 1999, 268)

péŋ

aʔ=wĩ-n

aket

white.man

3SG=kill-IND1

INFR

'The white man must have killed it.' (used in a situation where it was known by the speaker that the white man had gone in the forest overnight to hunt and came back with his prey, but neither the speaker nor anybody else saw him killing it)

(original gloss *aket* = must)

ASSUMED is the Evidential that is furthest removed from direct experience (of all the Evidential categories in the questionnaire) and refers to general knowledge. For example, in (6.9) the speaker assumes that the subject of the clause has gone to sleep because several circumstances point to that conclusion, including the knowledge that people go to sleep at night. This conclusion is drawn on the basis of general knowledge (it being late, people usually go to sleep). Compare this to the Karo marker *aket* in (6.8) above, where there is visual evidence of the result (the killed animal):

- (6.9) Karo (Tupían; Gabas Jr. 1999, 268)

to=ket-a

aʔ=eʔ-t

igã

3R=sleep-GER

3SG=AUX-IND1

ASM

'He must have gone to sleep.' (used when the speaker kept waiting for a person for a long time, it was late at night, and the person did not show up. So, the speaker concludes that the person may have gone to sleep.)

(original gloss *igã* = *igã*)

6.3 TYPOLOGICAL DISTRIBUTION IN THE SAMPLE

6.3.1 General overview

Of the 63 languages in the sample, 14 do not mark any Evidential, but 49 have at least one morpho-syntactically marked Evidential category. Of these 49 languages, 22 mark a direct witness (i.e. *FIRSTHAND* and/ or *VISUAL*), and 46 an indirect witness Evidential (i.e. *NON-FIRSTHAND*, and/ or *SECONDHAND*, and/ or *QUOTATIVE*, and/ or *INFERRED* and/ or *ASSUMED*). An intersecting set of 20 languages mark both direct and indirect witness. Split up into the individual categories, results show that the most frequently marked category is *SECONDHAND* (38), followed by *FIRSTHAND* (17), *INFERRED* (15), *NON-FIRSTHAND* (11), *ASSUMED* (10), *VISUAL* (5), and *QUOTATIVE* (2). 23 languages in the sample mark one Evidential category, 11 mark two, nine mark three, four mark four, and two mark five.

The form of marking ranges from suffixes over clitics to particles; there is one prefix (*SECONDHAND* *kidi-* in Itonama). Suffixation is most common in the sample. However, Evidentiality is marked relatively more frequently by particles than Tense, Aspect, and Modality. There are eight languages with cumulative morphemes of Evidentiality and Tense, usually involving *PAST* and sometimes *PRESENT*, rarely *FUTURE*. Languages marking Evidentiality and Tense with cumulative morphemes are: Tariana, Aymara, Jarawara, Mamaindê, Sabanê, Matses, Cubeo, and Desano. Aymara is the only language involving *FUTURE* in a cumulative morpheme with Evidentiality. De Haan (2011) reports the same distribution for a global sample: “[m]ost languages that use the verbal system to denote evidential distinctions do so only in the past tense”.

The following tables and figures illustrate the distribution of Evidentiality marking in the sample. Table 6.1 shows the number of languages marking Evidentiality or not. There are 49 languages marking direct and/ or indirect Evidentiality, 21 of which mark both. Furthermore, direct and indirect are split into their respective categories and how many languages mark these. Figure 6.2 shows how many languages mark which category and figure 6.3 how many languages mark which number of Evidentials. These numbers are all based on table 6.4, which presents the languages and which Evidential categories they mark/ do not mark, ordered from least to most marking.

The following section (6.3.2) presents the Evidentiality systems occurring in the sample. A detailed discussion of geographical and genealogical patterns follows in section 6.4 and 6.5.

Table 6.1: Schematic overview of Evidentiality marking in the sample

No Evid.	Evidential marking					
14	49					
	direct 22		indirect 46			both 21
	FIRSTHAND	VISUAL	NON-FIRSTHAND	SECONDHAND (+QUOTATIVE)	INFERRED	ASSUMED
	17	5	11	38(+2)	15	10

Figure 6.2: Number of SAILs marking number of Evidential categories

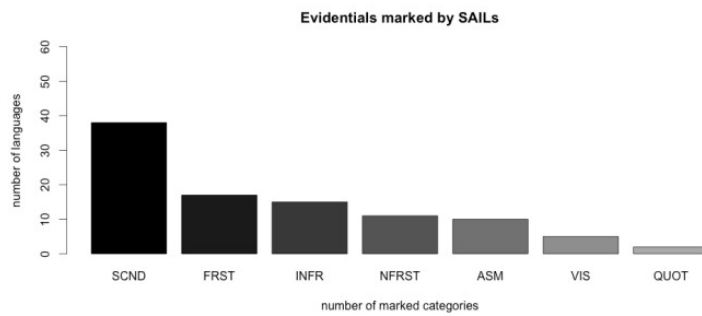
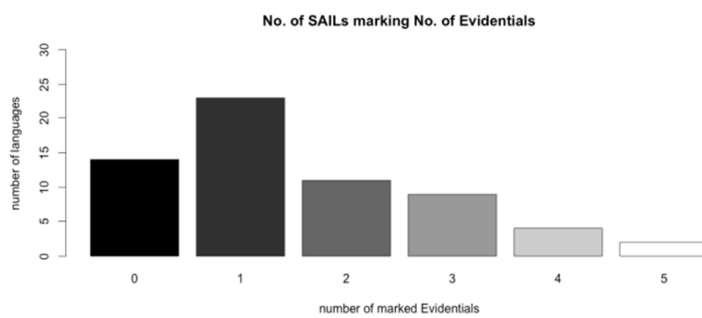


Figure 6.3: Evidential categories marked by SAILs



34	Kwaza	Unclassified								1
35	Mapuche	Araucanian								1
36	Mocoví	Guaycuruan								1
37	Mekens	Tupían								1
38	Urarina	Unclassified								2
39	Imbabura Quechua	Quechuan								2
40	Tapiete	Tupían, Guarani								2
41	Tsafiki	Barbacoan								2
42	Miraña	Boran								2
43	Sateré-Mawé	Tupían								2
44	Kamaiurá	Tupían, Guarani								2
45	Cocama- Cocamilla	Tupían, Guarani								2
46	Ika	Chibchan, Aruak								2
47	Huallaga Quechua	Quechuan								2
48	Yanam	Yanomaman								2
49	Hup	Nadahup								3
50	Mosetén	Mosetenan								3
51	Sabanê	Nambikwaran								3
52	Jarawara	Arawan								3
53	Desano	Tucanoan								3
54	Yurakaré	Unclassified								3
55	Leko	Unclassified								3
56	Matses	Panoan								3
57	Aymara	Aymaran								3
58	Shipibo- Konibo	Panoan								4
59	Cuzco Quechua	Quechuan								4
60	Cubeo	Tucanoan								4
61	Karo	Tupían								4
62	Tariana	Arawakan								5
63	Mamaindê	Nambikwaran								5
			38	17	15	11	10	5	2	

6.3.2 Systems of Evidentiality marking

This section presents an overview of the systems of Evidentiality marking as occurring in the sample. Note that the way the term ‘system’ is referred to here does not coincide with Aikhenvald’s (2003) more narrow definition. Aikhenvald includes formally unmarked Evidential functions, i.e. zero-markers. That leads to systems that can have unmarked members, in contrast to the present definition of a system, where all members must be overtly marked. The reason that I have chosen to include only marked categories is the fact that for many of the languages in the present sample it is not possible to establish obligatoriness, which was defined in chapter 2 as a necessary condition for the postulation of zero-morphemes. As a result, the present systems usually have one Evidential category less than in Aikhenvald’s (2004) analysis. Additionally, an Evidential system as defined here does not specify whether the markers within one language constitute a paradigmatic system, are in complementary distribution, or occupy different verbal slots. Aikhenvald (2003) makes a distinction between “unitary” versus “scattered” systems, i.e. those with a paradigmatic set of forms versus several Evidential forms distributed throughout the grammar. This will be commented upon when relevant.

As can be seen in table 6.4 there is a cline between the number of Evidentials marked and the number of languages marking it, e.g. more languages mark a small number of Evidential categories while fewer languages mark a higher number. The most frequent system of Evidentiality marking in the sample has only one option (mostly *SECONDHAND*). The questionnaire allows for a maximum of seven options, but the highest number of options is five in only two languages (Tariana and Mamaindê).

The following section presents the individual systems with illustrative examples. A full discussion of Evidential marking in each language is beyond the scope of this study; instead, illustrative and a few peculiar cases will be investigated.

Evidentiality systems with one option

This is the largest group in the sample, found in 23 languages. The majority of these languages show *SECONDHAND* marking (17), two exhibit *FIRSTHAND*, two *NON-FIRSTHAND*, one has *QUOTATIVE*, and one has *INFERRED*. A system with only one overtly marked option may seem like a binary one where one category is unmarked or zero-marked.

SECONDHAND is by far the most frequently marked category, and Aikhenvald (2004, 31-32) already showed that simple Evidentiality systems of the sort “reportative versus everything else” are indeed common throughout South America. All of the languages in this category have one or more markers which indicate that the information was acquired by verbal report, such as in Karitiana:

(6.11) Karitiana (Tupían; Everett 2006, 282)

a-taka-tat-sarít-∅

2SG.ABS-NSAP-go-SCND-NFUT

‘You went, I heard.’

Much less common in the sample are one-option systems with markers for FIRSTHAND, NONFIRSTHAND, INFERRED, or QUOTATIVE. There is no language which has an ASSUMED marker only. The following paragraphs investigate these subtypes.

Both Mocoví and Mapuche have a NON-FIRSTHAND marker. In the case of Mocoví, the enclitic =o? refers to unwitnessed events and is extremely common in narratives (Grondona 1998, 180). It is attached to the verb ‘say’ before a direct speech construction and thus also marks a QUOTATIVE. The NON-FIRSTHAND suffix -(ü)rke in Mapuche is also commonly used in narratives, and marks information by deduction, report, or general hearsay. An uncommon meaning among the NON-FIRSTHAND markers in the sample is that the Mapuche suffix is also used in contexts where the participant (not the speaker) was unaware of some aspect of the verbal action (Smeets 2008, 246f.) - see (6.13b).

(6.12) Mocoví (Guaycuruan; Grondona 1998, 212)

- (a) *ka?* *i+ɣyak+o?* *l+asom* *l+ai*
 then 3ACT+bring+NFRST 3POSS+door ABS+side
 ‘And she placed it near the door (=by the side of the door).’ (taken from a story told to the speaker by his grandfather)
- (b) *ka?* *∅+na:k+o?* *ka* *?yat*
 then 3ACT+say+NFRST DEIC.ABSN mosquito
 ‘Then the mosquito said: ...’

(6.13) Mapuche (Araucanian; Smeets 2008, 246, 247)

- (a) *amu-rke-lle-y-ng-ün*
 go-NFRST-AFM-IND-3NSG-PL
 ‘[oh yes,] they certainly went, I am told.’
- (b) *weñe-nie-ñma-rke-fi-y-ng-ün* *ñi* *mapu*
 steal-PRPS-IDR.OBJ-NFRST-EDO-IND-3NSG-PL POSS3 land
 ‘They kept robbing them of their land without them being aware of it.’

Mekens is the only language that has QUOTATIVE as the only marked Evidential category. QUOTATIVE is marked in two languages in the sample; usually SECONDHAND markers or complement clauses with verbs of the kind ‘say, tell, speak’ are used to express QUOTATIVE meaning. The Mekens QUOTATIVE particles *ke* (first and second person) and *kaat* (third person) predominantly occur under real quotative conditions, i.e. syntactically with a direct quote and an identified source, but also in more SECONDHAND-like environments. The QUOTATIVE marker *kaat* is interesting in another respect as well: It is possibly the second unit in the DESIDERATIVE *kot kaat* together with the IMMEDIATE FUTURE particle *kot*.

(6.14) Mekens (Tupían; Galucio 2001, 81, 207, 63)

- (a) *üsi nejat ikāō o-i-mi kaat*
 deer SML that.time 1SG-OBJ-kill QUOT
 ‘It looks like it is a deer that I shot that time’, he said.’
- (b) *po-kāra ese-ip pa òt kaat marly*
 hand-beads CMT-come FUT I QUOT Marly
 ‘Marly said that she would bring the bracelets.’ (lit. ‘Marly said: “I will bring the bracelets.”’)
- (c) *kaab=eri eba òt aose na eteet*
 that=ABL EV I person VBLZ could
o-iki ò-a ke te
 1SG-water give-THEM QUOT truly
 ‘That is why I said at that time that if it were a person it should give me water.’

Kwaza is the only language which has an INFERRED marker in a one-option system, and is worth further consideration. There is little doubt that the suffix *-tehere* (and *-cehere*, which is probably its allomorph) marks INFERRED Evidentiality (Van der Voort 2004, 420). In example (6.15a), the sentence not marked with *-tehere* is neutral about the source of information, but in example (6.15b), *-tehere* clearly marks that the proposition is based on visible evidence (the ground being wet):

(6.15) Kwaza (unclassified; Van der Voort 2004, 421)

- (a) *a'we-hỹ-ki*
 rain-NOM-DECL
 ‘It has rained.’ (I did or did not see it rain.)
- (b) *a'we-hỹ-tehere*
 rain-NOM-INFR
 ‘It must have rained.’ (I did not see it rain, but the ground is wet.)

The suffix *-tehere* is the only morpho-syntactic Evidential marker in Kwaza. Only one other marker in Kwaza has any Evidential meaning: the indefinite subject suffix *-wa*, which marks unspecified third person(s). It can also indicate direct or indirect witness, but this is considered a “side-effect” by Van der Voort (2004, 264) and corresponds to what Aikhenvald (2004, 105) calls an “evidential strategy”, which includes any marker that acquired secondary Evidential meanings in contrast to markers with primary meaning. As is the nature of grammaticalization, a former marker with secondary Evidential meanings may develop into a full Evidential, and the Kwaza suffix *-wa* may be in such an early stage of grammaticalization.

It is likely that *-tehere* has developed language-internally rather than having been borrowed. The form *-here* occurs in three different, but apparently related morphemes which Van der Voort (2004, 417-418) summarizes as “conjunctural”: the INTENTIONAL marker

-here, which probably developed from the negation suffix *-he* and the interrogation suffix *-re*, the appellative marker *-herejã*, and the INFERRED marker *-tehere*. Given the fact that the INTENTIONAL marker can be historically separated into two different morphemes, and that the Evidential and appellative morphemes have an additional element, it is possible that the latter two derive from the INTENTIONAL. Such a grammaticalization path is yet to be proven, but seems the most likely explanation for the existence of the INFERRED marker in Kwaza.

Evidentiality systems with two options

Evidentiality systems with two Evidential categories marked in the sample can be grouped into systems with a) one direct and one indirect witness marker, and b) two indirect witness markers. The following section discusses these two subsystems and their internal division.

The first subsystem consists of languages that mark both direct and indirect witness with one marked category each. This group can be further broken down into FIRSTHAND/ NON-FIRSTHAND (Ika, Yanam, Tapiete), FIRSTHAND/ SECONDHAND (Urarina, Huallaga Quechua, Imbabura Quechua), and VISUAL/ SECONDHAND (Kamaiurá, Nheengatú) marking.

Ika, Yanam, Tapiete, and Huallaga Quechua all have a FIRSTHAND/ NON-FIRSTHAND option, which appears to be obligatory in none of the languages. A speaker of Huallaga Quechua can choose between the FIRSTHAND suffix *-mi* or the NON-FIRSTHAND suffix *-shi*, or neither. In Ika, the FIRSTHAND suffix *-in* seems to be restricted to visual evidence (according to the examples given), but because Frank (1985, 105) labels it direct witness without further comment it is taken as FIRSTHAND. Its Evidential opposite is the particle *ni*, as can be seen in (6.16):

(6.16) Ika (Chibchan, Aruak; Frank, 1985, 105)

- | | | | |
|-----|------------------------------------|----------------|------------|
| (a) | <i>win-naka</i> | <i>u-ž-in.</i> | |
| | 3PL.SBJ-COME | AUX-MED-FRST | |
| | 'They came (and I saw it).' | | |
| (b) | <i>win-naka</i> | <i>u-na</i> | <i>ni.</i> |
| | 3PL.SBJ-COME | AUX-DIST | NFRST |
| | 'They came (but I didn't see it).' | | |

The two subgroups FIRSTHAND/ SECONDHAND and VISUAL/ SECONDHAND are represented by Urarina, Huallaga Quechua, and Imbabura Quechua for the former, and Kamaiurá and Nheengatú for the latter. Urarina distinguishes between direct and indirect witness with a FIRSTHAND clitic =*ni* and two SECONDHAND markers. They do not constitute a binary Evidential system as their usage is entirely optional. The FIRSTHAND marker is most common as an eyewitness marker, but can also refer to any other sensory evidence. The two SECONDHAND markers are the enclitic =*he* for information that is acquired through verbal report from a third person, and a particle, *hetau*, which occurs in narratives, especially those with distant past reference (Olawsky 2006, 496ff., 862).

- (6.17) Urarina (unclassified; Olawsky 2006, 497, 499, 918)
- (a) *it̥sa-i=ni=tau*
do-2=FRST=REAS
'You did it earlier today.' (I know it, because I was a witness.)
- (b) *rauto-a=ni*
be.tasty-3=FRST
'It was tasty.' (Referring to what I ate in the morning.)
- (c) *it̥safwa-a=ni=tau*
shoot-3=FRST=REAS
'He shot it (earlier today).' (I saw or heard him shooting.)
- (d) *it̥sa-k̥a=he*
do-1-SCND
'He says that I did it.'
- (e) *nii h̥a h̥etau=te tem̥-a k̥a-ure*
that because SCND=FOC plant-NTR go-3PL
'Therefore, they went to plant.' (from a traditional narrative about the creation of the peccaries)

For Huallaga and Imbabura Quechua, see the Quechuan section below.

The second subsystem consists of languages with an option of two Evidentials marking two different indirect witness Evidentials. They all mark SECONDHAND: Tsafiki marks SECONDHAND and INFERRED, Miraña marks NON-FIRSTHAND and SECONDHAND, and Cocama-Cocamilla marks SECONDHAND and QUOTATIVE. Cocama-Cocamilla is one of the two languages in the sample that explicitly marks a specified source with a direct speech construction (the other is Mekens), and does so in addition to having a SECONDHAND marker. The SECONDHAND marker =*ía* refers to an unspecified source, whereas the QUOTATIVE particles *ay* and *na* are predominantly used in constructions with a direct quote and a verb of telling. They differ in that *ay* introduces a quote, whereas *na* follows it. The particle *na* can also occur without a direct quote (Vallejos Yopán 2010, 172).

- (6.18) Cocama-Cocamilla (Tupán; Vallejos Yopán 2010, 495, 173, 174)
- (a) *manir=ía aykua-pa*
Manuel=SCND be.sick-COMPL
'It is said that Manuel is very ill.'
- (b) *na piata tsewe, na ra kumitsa*
2SG ask salt QUOT 3SM say
'"You ask for salt," he says.'
- (c) *uri, eee, kumitsa ay na papa tsapuki-ta=ene*
3SG.LM hm says QUOT 2SG father call-CAU=2SG.L
'He says "Your father calls you".'

Tsafiki has *SECONDHAND* and *INFERRED* markers, but also an unmarked *FIRSTHAND* which seems to be in direct opposition to the other markers (Dickinson 2002, 102). Because Dickinson does not comment on whether Evidential marking in Tsafiki is obligatory or not, it cannot be said whether this is a zero-morpheme.

Evidentiality systems with three options

Of the languages in the sample that offer an option of markers for three Evidential categories seven mark both direct and indirect witness and two only indirect witness. All of the former have *FIRSTHAND* as direct witness category and can thus be grouped into a *FIRSTHAND/indirect witness* opposition. They are structured into *FIRSTHAND/ NON-FIRSTHAND/ SECONDHAND* (Jarawara), *FIRSTHAND/ NON-FIRSTHAND/ INFERRED* (Aymara), *FIRSTHAND/ SECONDHAND/ INFERRED* (Sabanê, Hup, Mosestén), *FIRSTHAND/ SECONDHAND/ ASSUMED* (Desano), and *FIRSTHAND/ INFERRED/ ASSUMED* (Matses). The two languages that mark indirect witness only in an Evidentiality system of three options have the same distribution of markers: *SECONDHAND/ INFERRED/ ASSUMED* (Yurakaré, Leko). The following section gives an overview of Hup as a representative of a three-option system with a direct/ indirect witness opposition.

Hup distinguishes between three optionally marked Evidential categories, and a possibly zero-marked fourth one. Although Evidentiality marking is not obligatory in Hup, the absence of *FIRSTHAND* marking is usually understood as coding *VISUAL* (Epps 2008, 641ff.). Usage of the *FIRSTHAND* marker is restricted to sensory but non-visual experience. This is unusual in the sample as most *FIRSTHAND* markers include vision.

(6.19) Hup (Nadahup; Epps 2008, 646)

- (a) *náciya pæ-cíw-íy=hã*
 boat go.upriver-COMPL-DYNM=FRST
 ‘The boat already went upriver.’ (speaker heard but did not see it)
- (b) *?ok-níh key-ham-g’et-yí?-ay=∅ tíh=?áy-áh!*
 move=NEG see-go-stand-TEL-INCH=VIS 3SG=FEM-DECL
 ‘She was just standing there looking, without moving!’ (speaker witnessed event)

Fleck (2003, 399ff.) makes a special point of the fact that the Evidential system in Matses (which is obligatory and fused with *PAST* Tense) makes a distinction between whether the event was witnessed at the same time as it occurred, i.e. simultaneously, or not. The former is marked by *FIRSTHAND PAST* suffixes, whereas the latter are marked by *INFERRED PAST* and *ASSUMED PAST* suffixes. Logically, this is the case with all Evidentials cross-linguistically; an event can only be directly witnessed when the witness is there the same time the event occurs.

Evidentiality systems with four options

All of the four languages with four options of marked Evidential categories have one direct-witness category and three indirect ones: *FIRSTHAND*, *SECONDHAND*, *INFERRED*, and *ASSUMED*

(Shipibo-Konibo, Cubeo, Cuzco Quechua), and VISUAL, SECONDHAND, INFERRED, and ASSUMED (Karo).

Shipibo-Konibo has markers for FIRSTHAND, SECONDHAND, INFERRED, and ASSUMED Evidentiality. The corresponding forms *ri*, *ronki/ki*, and *bira* occur most frequently in enclitic (but also in proclitic) position and are not verbal morphemes *per se* but can attach to various clause constituents. Evidentiality marking in Shipibo-Konibo is also obligatory “in the sense that the evidential value of the information has always been grammatically marked in the forgoing discourse and is clear to native speakers” (Valenzuela 2003a, 39). The FIRSTHAND marker =*ri* covers all sensory experiences and also marks the speaker as active participant, e.g. when telling a fact about himself. There are two markers for SECONDHAND Evidentiality, =*ronki* and =*ki*, with =*ki* probably being the shorter form of =*ronki*, which marks an identified or unidentified source of verbal report. The ambiguous clitic =*bira* marks ASSUMED and INFERRED (Valenzuela 2003b, 534ff.). This clitic is not a NON-FIRSTHAND marker because it does not include reportative or hearsay under its meanings.

(6.20) Shipibo-Konibo (Panoan; Valenzuela 2003, 536, 541, 545)

- (a) *e-a r-iki Bawanixo-nko-ni-a.*
 1-ABS FRST-COP Bawanixo-LOC-LIG-ABL
 ‘I am from Bawanisho.’
- (b) *a-ronki-ai.*
 do-SCND-INCP
 ‘It is said that she will do (it)./ She says that she will do it.’
- (c) *beka-ra pi-kas-i-bira i-t-ai.*
 Beka:ABS-FRST eat-DES-SSSS-INFR be-PROG-INCP
 ‘Beka must be hungry.’ (Since everybody knows there is no fish in her village)

In contrast to the direct witness markers in Shipibo-Konibo and Cubeo, the one in Karo refers only to eye-witness. The VISUAL marker *topə* marks events that were eye-witnessed by the speaker, but can also refer to a noun phrase which is the subject of the witnessed event. In example (6.21b), *topə* provides “information about the head noun and the proposition” (Gabas Jr. 1999, 272). Furthermore, Karo marks SECONDHAND with the particle *tə*, INFERRED with *aket* and ASSUMED with *igā*³ and *memā* (ibid. 266 ff.).

(6.21) Karo (Tupían; Gabas Jr. 1999, 266, 273, 267, 268)

- (a) *péj ?e-t topə to=at maca?it wĩ-a*
 white.man AUX-IND VIS 3=POSS pet kill-GER
 (It was seen that) the white man killed his pet.’ (used in a situation where the speaker went to the white man’s house and saw him killing his pet.)

³ This marker is formally similar to the FUTURE particle *iga* (cf. Gabas Jr. 1999, 183).

- (b) *péŋ* **topə** *toat* *macaʔit* *nō* *wī-n*
 white.man VIS 3.POSS pet one.of kill-IND
 ‘The white man was seen to have killed one of his own pets.’
- (c) *péŋ* *aʔ=wī-n* **aket**
 white.man 3SG=kill-IND INFR
 ‘The white man must have killed it. (used in a situation where it was known by the speaker that the white man had gone in the forest overnight to hunt and came back with his prey, but neither the speaker nor anybody else saw him killing it)
- (d) *to=ket-a* *aʔ=ʔe-t* **igã**
 3r=sleep-GER 3SG=AUX-IND ASM
 ‘He must have gone to sleep.’ (used when the speaker kept waiting for a person for a long time, it was late at night, and the person did not show up. So, the speaker concludes that the person might have gone to sleep)
- (e) *aʔ=ket-t* **memã**
 3SG=sleep-IND ASM
 ‘I suppose he is sleeping.’ (used in a situation where the speaker knew that the person in question was sleeping before)

Evidentiality systems with five options

The two languages that mark five Evidential categories in the sample have the same semantic distribution: both Tariana and Mamaindê mark VISUAL, NON-FIRSTHAND, SECONDHAND, INFERRED, and ASSUMED Evidentiality. These are the most elaborate Evidential systems in the sample. It is notable that the markers have a high a degree of fusion, i.e. almost every single marked Evidential category is fused with at least one other category. The exception is the Mamaindê SECONDHAND suffix *-satau* which has no other, non-Evidential meaning. All of the markers are suffixes (Mamaindê) or enclitics (Tariana). When there is a cumulative suffix for a category, Evidentiality is always fused with Tense. Mamaindê has one set of ambiguous (and cumulative) markers for INFERRED and ASSUMED, i.e. the same markers are used to express either of them. Additionally, Mamaindê marks ASSUMED Evidentiality that refers to general knowledge as the information source, with a separate set of suffixes.

Mamaindê is the only language in the sample that has a THIRDHAND marker, i.e. which specifically marks that the verbal report given to the speaker was also acquired by a verbal report. This marker, the suffix *-sîn*, contrasts with the SECONDHAND suffix *-satau* which denotes that the verbal report was given to the speaker but does not further specify whether it originated in yet another person (Eberhard 2009, 460ff.). In Tariana, the SECONDHAND markers cover both SECONDHAND, THIRDHAND, and more removed reports (Aikhenvald 2003, 302).

(6.22) Mamaindê (Namikwaran; Eberhard 2009, 469)

- (a) *waʔnĩn-soʔka* *jajãn-tu* *sun-satau-le-ø-hĩn-wa*
 shaman-NCL.HUM jaguar-FNS kill-SCND-PST-SBJ3-PST/NVIS-DECL
 ‘The shaman killed a jaguar (yesterday).’ (and I know this because someone told me)
- (b) *waʔnĩn-soʔka* *jajãn-tu* *sun-sĩn-ø-nha-wa*
 shaman-NCL.HUM jaguar-FNS kill-THRD-SBJ3-PRS/NVIS-DECL
 ‘The shaman killed a jaguar (yesterday).’ (and I know this because someone said they were told that it was so)

The VISUAL Evidentiality markers in both Mamaindê and Tariana have a meaning extension in the direction of gnomic statements: the VISUAL in Tariana not only refers to eyewitness accounts, but also to general and observable facts (Aikhenvald 2003, 295). In Mamaindê, there is an extra set of markers for general knowledge, but the VISUAL markers are also used as default for obvious facts (Eberhard 2009, 463-464).

(6.23) Tariana (Arawak; Aikhenvald 2003, 296)

- ikasu* *yaphini-nuku* *itʃiri* *ma-inu-kade-naka* *nemhani*
 today thing;like-TOP.NON.A/S game NEG-kill-EG-PRES.VIS 3PL+go.round
 ‘On a day like today no one goes around hunting.’

Table 6.3 provides an overview of the Evidential systems as occurring in the sample, ordered from one to five options.

Table 6.3: Evidentiality systems

Number of lgs. marking option x		Number of languages		Number of languages	Language(s)
One option (23)	direct	2	FIRSTHAND	2	Emérillon, Imbabura Quechua
	indirect	21	SECONDHAND	17	Kaingang, Nasa Yuwe, Yanesha', Trumai, Karitiana, Warao, Itonama, Bororo, Baure, Apurinã, Yaminahua, Cavineña, Puinave, Hixkaryana, Movima, Embera, Dâw
			NON-FIRSTHAND	2	Mocoví, Mapuche
			QUOTATIVE	1	Mekens
			INFERRED	1	Kwaza
Two options (11)	direct/ indirect	7	FIRSTHAND/ NON-FIRSTHAND	3	Ika, Yanam, Tapiete
			FIRSTHAND/ SECONDHAND	2	Urarina, Hualla Quechua, Imbabura Quechua
			SECONDHAND/ VISUAL	2	Kamaiurá, Nheengatú
	indirect/ indirect	3	SECONDHAND/ INFERRED	1	Tsafiki
			NON-FIRSTHAND/ SECONDHAND	1	Miraña
			SECONDHAND/ QUOTATIVE	1	Cocama-Cocamilla
Three options (9)	direct/ indirect/ indirect	7	FIRSTHAND/ NON-FIRSTHAND/ SECONDHAND	1	Jarawara
			FIRSTHAND/ NON-FIRSTHAND/ INFERRED	1	Aymara
			FIRSTHAND/ SECONDHAND/ INFERRED	3	Sabanê, Hup, Mosestén
			FIRSTHAND/ SECONDHAND/	1	Desano

			ASSUMED		
			FIRSTHAND/ INFERRED/ ASSUMED	1	Matses
	indirect/ indirect/ indirect	2	SECONDHAND/ INFERRED/ ASSUMED	2	Yurakaré, Leko
Four options (4)	direct/ indirect/ indirect/ indirect	4	FIRSTHAND/ SECONDHAND/ INFERRED/ ASSUMED	3	Cubeo, Shipibo-Konibo, Cuzco Quechua
			VISUAL SECONDHAND/ INFERRED/ ASSUMED	1	Karo
Five options (2)	direct/ indirect/ indirect/ indirect/ indirect	2	VISUAL/ NON-FIRSTHAND/ SECONDHAND/ INFERRED/ ASSUMED	2	Tariana, Mamaindê

6.4 GEOGRAPHICAL DISTRIBUTION

6.4.1 Results

The geographical distribution of morpho-syntactic Evidential marking in the sample is very heterogeneous. This section discusses the similarities and differences of marking within families and cross-linguistically with regard to possible regional clusters. It attempts to give a broad overview, with individual cases being discussed in more detail. Where possible, comparisons to additional studies of Evidentiality marking, globally or more specific will be given.

There are 22 languages marking direct and 46 marking indirect Evidentiality; 20 languages mark both. Only one language in the sample marks direct Evidentiality only: Emérillon. Its *FIRSTHAND* suffix *-lako* codes direct witness but also past time reference, and is quite rare (Rose 2003, 440). Rose (ibid. 441) keeps the possibility open that *-lako* is rather an Epistemic Modal marker. If this turns out to be true, it would solve the problem that Emérillon is the only language in the sample that exclusively marks direct Evidentiality. Nevertheless, the examples in Rose (2003) are not extensive enough for a more detailed analysis. Additionally, it is much more likely that *-lako* is primarily a *PAST* marker.

(6.24) Emérillon (Tupían, Guaraní; Rose 2003, 440)

aj kʷalai-l-che-lako Roger o-singal a Niwe.
 DEM year-RELN-POSTP-PST.FRST Roger 3-special.form.of.fishing at Niwé
 ‘In that year, Roger fished at Niwé (and I was there).’ (orig. ‘Cette année-lá, Roger a
 nivré à Niwe (et j’y étais.)’)

Emérillon is an exception not only in this sample, but world-wide: In a global study of 418 languages De Haan (2011) does not find a single instance of a language only marking direct, without marking indirect Evidentiality as well. He reports 181 languages without grammaticalized Evidentiality, 166 with only indirect, and 71 with both kinds. In this sample the ratio of indirect and direct+indirect marking is more balanced: 27 mark indirect, 21 both, 14 none, and one direct Evidentiality. A study by Torres Sánchez (2008) of Evidentiality in 204 languages in North, Meso-, and South America (92 of these in SA), reveals also only one single language, Oaxaca (Tequistlatecan), that marks only direct, but not indirect Evidentiality.

The overall number of SAILs with morpho-syntactic Evidentiality marking is quite high (49 out of 63). Regarding the recent literature on Evidentiality in SA this is not surprising, and was already demonstrated by De Haan (2011) who found 237 languages with and 181 without Evidential markers world-wide; this translates for SA into 52 with and 25 without Evidentials. Torres Sánchez (2008) finds 119 languages with and 85 without Evidentials. All in all, the ratio of languages with Evidentials in this study is slightly higher than in De Haan and Torres Sánchez, but that may be due to the fact that the sample favors areas where Evidentiality reportedly occurs quite frequently (e.g. Amazonia (cf. Aikhenvald (2004), the Guaporé-Mamoré (cf. Crevels & Van der Voort (2008), and the Andes (cf. Adelaar 2012)). These are indeed areas where many Evidentials occur in the sample, but one can see on map 6.1 that Evidential marking occurs everywhere in SA and does not favor or neglect a certain greater region.

The frequent occurrence of Evidentials in general shows that it is probably a characteristic of, although not singular to, South American languages. De Haan (2011) and Torres Sánchez (2008) have shown that Evidential marking is frequent in all of the Americas as opposed to, for example, Africa (De Haan 2011). As has been demonstrated for certain regions in SA this is partially due to areal diffusion conditioned by the socio-linguistic settings, such as linguistic exogamy (e.g. Aikhenvald 1999, 2004; Epps 2005), but much more work needs to be done to account for the presence of Evidentiality in all of SA. Whether for the most part areal diffusion is responsible or indeed inheritance is an open question. As of now, it seems possible that Evidentiality developed in different areas independently from other areas, and that the aggregation of linguistic areas together forms the uniform picture of Evidentiality today. It would follow that Evidentiality is prone to diffusion and that languages without Evidentiality are then predominantly found outside linguistic areas. However, in order to test this hypothesis a large study of much more languages than the present sample gives us is necessary.

Table 6.4 compares two studies of Evidentiality world-wide (De Haan 2011) and in the Americas (Torres Sánchez 2008) with the present study. Unfortunately, it was not possible to establish the number of Evidentials in SA in Torres Sánchez. Map 6.1 illustrates (i) the distribution of languages marking and not marking Evidentiality morpho-syntactically and (ii) which languages mark direct or indirect witness, or both, or none. Map 6.2 illustrates the distribution of languages with either direct or indirect Evidentiality marking, or both or none of these.

Table 6.4: Evidentiality marking in three samples

	De Haan (2011) (global)	Torres Sánchez (2008) (North/ South America)	This study (South America)
Total number of lgs. globally	418	204	63
Total number of lgs. in SA	77	92	63
Evidentials present in complete sample	237	119	49
Evidentials present in SA	52	?	49

According to map 6.3, two languages that mark a large number of Evidentials occur in the same area, the Vaupés: Tariana (5 Evidential categories) and Cubeo (4 Evidential categories). De Haan (2011) reports a cluster of the largest Evidential systems in the Vaupés and adjacent areas well. The languages with the most Evidential markers in this sample are Tariana and Mamaindê. Whereas the former is indeed situated in the Vaupés, the latter is situated in another region entirely: west of the Guaporé-Mamoré area. Although Mamaindê and Tariana share the feature of a rich Evidential system (but are in different families), their Evidential systems developed for different reasons. The Tariana system largely results from language contact with Tucanoan (see below and Aikhenvald (2004)), but the Mamaindê system is largely genealogically conditioned (cf. Eberhard 2009, 451ff.). Although the two Nambikwaran languages in the sample, Mamaindê and Sabanê, have very different SECONDHAND markers (suffix *-satau* and suffix *-tiaka*, respectively), the Mamaindê suffix is formally and functionally similar to the Lakondê (also Nambikwaran) SECONDHAND markers *-'se?* and *-setaw* (Telles & Wetzels 2006, 239). Lakondê makes a distinction in marking a verbal report whose source is identified (*-setaw*) and not identified (*-'se?*); thus, the former has one characteristic of a QUOTATIVE, although it does not occur with direct quotation syntax. The Mamaindê suffix *-satau* on the other hand can be used with direct quotation although it typically does not identify the source (Eberhard 2009, 414). Nevertheless, this is evidence for a relationship of the SECONDHAND suffixes in Mamaindê and Lakondê.

There are a few peculiarities concerning the geographical distribution of the individual Evidential categories. One can observe that many of the languages marking ASSUMED also mark INFERRED Evidentiality. This is due to the fact that those languages have cumulative markers for both categories. On the other hand, there rarely are overlaps for SECONDHAND and QUOTATIVE, and VISUAL and FIRSTHAND. In many languages with a SECONDHAND marker this can also express QUOTATIVE, i.e. has a secondary QUOTATIVE meaning. From the two languages

with a QUOTATIVE marker, Mekens and Cocama-Cocamilla, Mekens does not have an additional SECONDHAND marker, but Cocama-Cocamilla does. In the latter case, there is no cumulative morpheme, but two different markers for QUOTATIVE and SECONDHAND.

Just like SECONDHAND often includes a QUOTATIVE meaning, FIRSHAND often includes VISUAL. Languages marking FIRSHAND rarely mark an additional VISUAL. It is rather the case that when FIRSHAND or VISUAL are marked, the other category is not marked. VISUAL marking languages are Tariana, Mamaindê, Karo, Sateré-Mawé, and Kamaiurá and none of these have an additional FIRSHAND marker or a cumulative morpheme for VISUAL and FIRSHAND. It is noteworthy that three of those are Tupían, which points toward a Tupían characteristic.

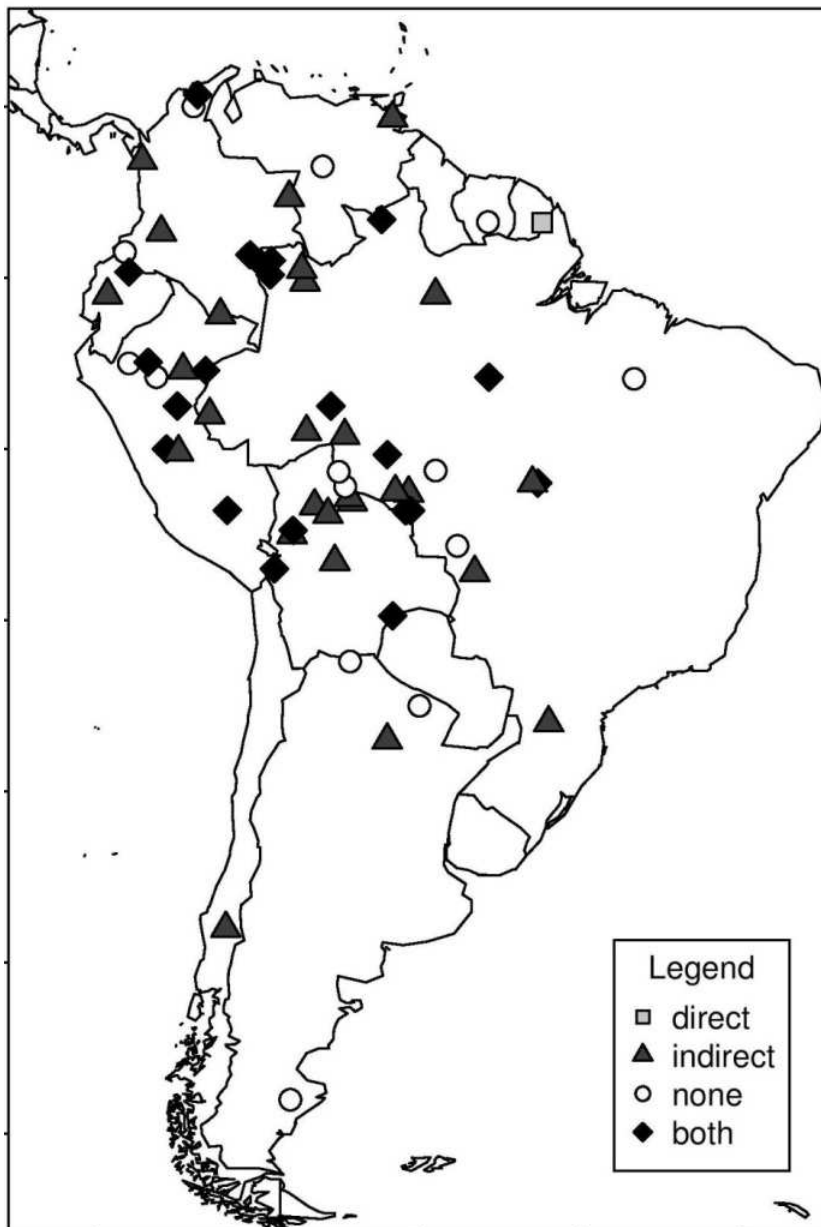
Map 6.1: Distribution of Evidentiality marking in the sample

Evidentiality marking



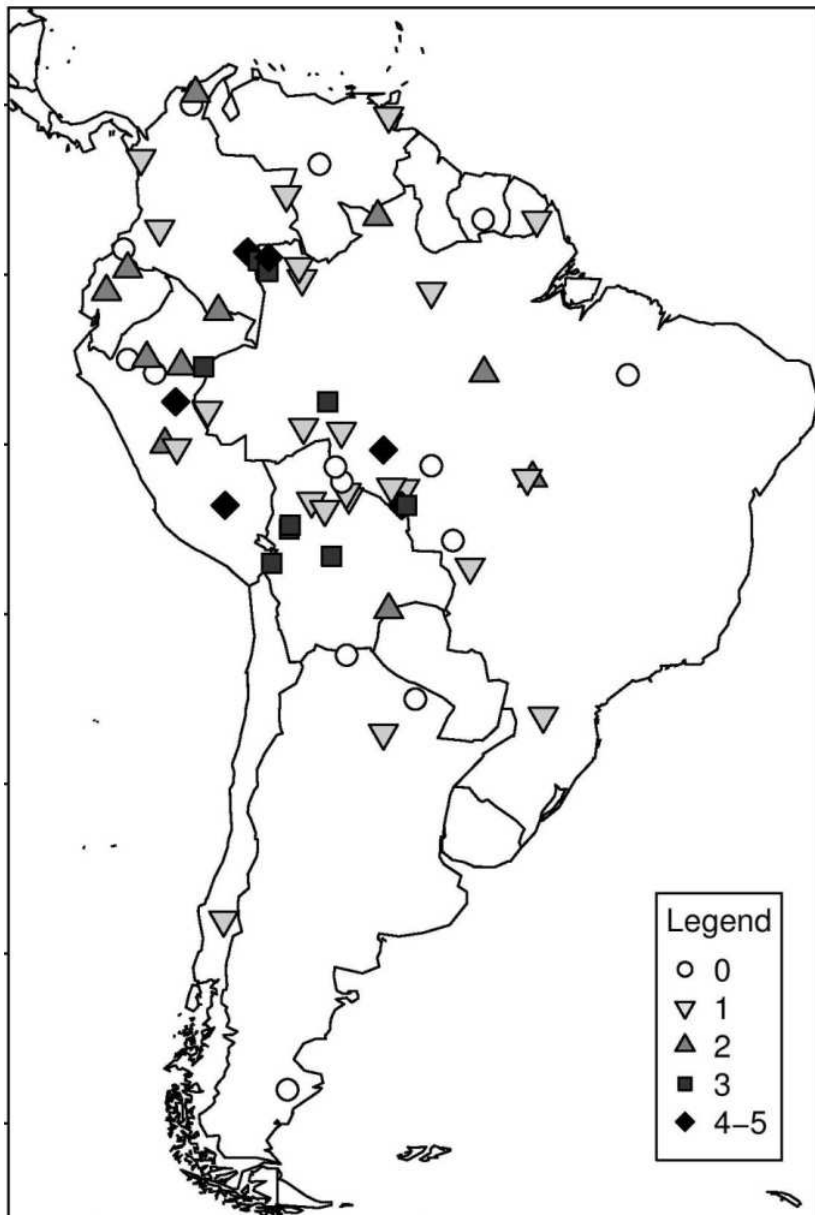
Map 6.2: Distribution of direct and indirect Evidentiality marking in the sample

Direct vs. Indirect Evidentiality



Map 6.3: Distribution of Evidential categories (by number) marked in the sample

Number of marked Evidentials



6.4.2 The Guaporé-Mamoré area

Crevels & Van der Voort (2008) propose that the Guaporé-Mamoré region in Rondônia, Brazil, and the Bolivian Amazonian lowlands constitutes a linguistic area. They give a provisional list of 55 languages present in that area largely based on geographical considerations (ibid. 166). Because Evidentiality occurs frequently in a subsample of 24 languages in the sample (approximately in 79,2%), and because Evidentiality has not been termed as “typically Amazonian”, they consider it a salient feature for a linguistic area (ibid. 170). However, Rodrigues (1999, 119) already mentioned that “Evidentiality is a pervasive feature of parts of Amazonia”, although he did not specify this for the Guaporé-Mamoré region. Aikhenvald (2004, 292, 303) similarly presents evidence for a cluster of grammaticalized Evidentiality in Amazonia. This rather points towards Evidentiality not just being a characteristic confined to the Guaporé-Mamoré, but for entire Amazonia (and also the Andes). The following paragraphs investigate whether the argument that Evidentiality marking constitutes a salient feature of the Guaporé-Mamoré linguistic area can be confirmed.

Crevels & Van der Voort’s (2008) present a list of 55 languages present in the Guaporé-Mamoré and further a detailed study of several proposed features in 24 languages for that area. Of these 24 languages, the present sample includes 13: Baure, Mosestén, Kwaza, Itonama, Leko, Movima, Mekens, Kanoê, Wari’, Cavineña, Yurakaré, Karo, and Karitiana.⁴ 11 of these show morpho-syntactic Evidentiality marking. The most frequently marked category is *SECONDHAND* (9), which is not surprising because it is the most frequently marked Evidential in the complete sample, followed by *INFERRED* (5), *ASSUMED* (3), *VISUAL* and *FIRSTHAND* (1 each), and *QUOTATIVE* (1). No language marks *NON-FIRSTHAND*. The form of marking ranges from suffixation to enclitics over particles to one prefix (Itonama *SECONDHAND*). The samples largely agree on which language marks Evidentiality, although in the present sample Wari’ does not and Yurakaré does. This may be due to different coding strategies.

There is almost no formal similarity between markers except for the *INFERRED* morphemes in Karo and Mosestén (see table 6.8 below). This leads to the conclusion that if the widespread Evidential marking is due to language contact, it is not the result of direct borrowing of form (unless the markers have undergone change processes that eradicated all formal similarities). It is more likely that they are the result of language internal processes, which nevertheless may have been triggered by contact. This would be evidence that Evidentiality in the Guaporé-Mamoré region is indeed a valid factor of a linguistic area.

Karo and Mosestén have formally and functionally similar *INFERRED* markers: the suffix *-(a)ke* and the particle *aket*, respectively. The Mosestén suffix *-(a)ke* is unproductive and “only appears in place names and other lexicalized forms” (Sakel 2004, 341), whereas the Karo particle *aket* seems to be fully productive. Karo (Tupían) and Mosestén (Mosestenan) are neither related nor geographically adjacent or even close. Together with Chimane and

⁴ I disregard Quechua, Uru, and Aymara as Crevels & Van der Voort (2008, 170) added them to the sample as control languages.

Mosetén de Santa Ana, Mosetén forms the small language family Mosetenan (Sakel 2004, 1-2). Chimane has the suffix *-aque* for events the speaker did not see or was not present at (Gill 1999, 124). Geographically close to Mosetén are Yurakaré and Leko, but both do not have a formally similar marker. In Yurakaré, INFERRED Evidentiality is marked by the suffix *-tiba*, and in Leko by *-nem*. Geographically, Karo and Mosetén are at the outskirts of the proposed linguistic Guaporé-Mamoré area; Karo is at the northeastern side and Mosetén at the southwestern side. Thus, with regard to the facts that Karo and Mosetén are not genealogically related and that speakers (at least today) are situated far away from each other and that none of the languages in between show formally similar markers at this point, it is most likely that they are the result of independent development. Furthermore, the form *aket* is not found as INFERRED marker in the other Tupían languages in the sample. Further research about possible historical contact between Karo and Mosetén speakers is necessary to see whether the INFERRED markers are due to language contact after all.

In several languages forms similar to the INFERRED markers in Karo and Mosetén are found that, however, do not express Evidentiality: Movima: IRREALIS *ak* (Haude 2006, 438); Cavineña perlocative *-eke* (Guillaume 2008, 611), and *-aki/-ki* 'being typical' (ibid. 330); Mekens demonstrative *-eke* (Galucio 2001, 47); Yuki demonstrative *-ake* (Villafañe 2004, 206); Yaminahua *-ake* 'around' (Faust 2002, 136-137); and Pilagá DESIDERATIVE *-ake* (Vidal 2001, 282).

The markers in Mekens and Yuki (both Guaraní) are probably Tupían cognates, but for the rest of the languages no obvious relation to each other or to the INFERRED markers can be established.

The quantitative analysis of Evidentiality marking supports Crevels & Van der Voort's claim, as 11 out of 13 languages do mark at least one Evidential category. On the other hand, because the overall frequency of Evidentiality marking in SA is quite high, this does not constitute a feature specific to the Guaporé-Mamoré. But because this sample is too small to be significant, one would need a full analysis of Evidentiality in all 55 languages of the original list to make a valid statement.

Table 6.5: Comparison of Evidentiality marking in the present and the Crevels & Van der Voort (2008) subsample

Language	Family	evi in the present sample	evi in Crevels & Van der Voort 's (2008) list of 24 languages
Baure	Arawakan	SCND = <i>hi</i>	yes
Wari'	Chapacuran	no	yes
Mosetén	Mosetenan	FRST <i>ishtyi'</i> , SCND <i>katyi</i> , INFR -(<i>a</i>) <i>ke</i>	yes
Cavineña	Tacanan	SCND = <i>pa</i>	yes
Karitiana	Tupían	SCND - <i>sarit</i>	yes
Mekens	Tupían	QUOT <i>ke, koot</i>	yes
Karo	Tupían	SCND <i>tə</i> , VIS <i>topə</i> , INFR <i>aket</i> , ASM <i>igā</i> , <i>memā</i>	yes
Yurakaré	Unclassified	SCND = <i>ya</i> , INFR/ ASM = <i>tiba</i> , = <i>jtě</i>	no
Leko	Unclassified	SCND - <i>mono</i> , INFR - <i>nem</i> (- <i>tson</i>), ASM - <i>nem</i>	yes
Kwaza	Unclassified	INFR - <i>tehere/ -cehere</i>	yes
Movima	Unclassified	SCND <i>ʔan</i>	yes
Kanoê	Unclassified	no	no
Itonama	Unclassified	SCND <i>kidi-</i>	yes

6.5 GENEALOGICAL DISTRIBUTION

6.5.1 Introduction

I now turn from a geographic focus towards discussing the distribution of Evidentiality in language families. There is some homogeneity of Evidentiality marking within (but also across) families, although - possibly due to the fact that the sample is not very dense for any family - mostly there is no clear pattern. For example, in this sample two of the three Cariban languages (Tiriyó and Panare) do not mark any Evidentiality, whereas the third, Hixkaryana, marks SECONDHAND. The two Barbacoan languages also differ: Tsafiki marks SECONDHAND and INFERRED, but Awa Pit marks none. Two of the Macro-Gêan languages do not mark any Evidentiality (Rikbaktsa and Timbira), but both Bororo and Kaingang mark SECONDHAND. With regard to the fact that most of the Evidentiality marking is predominantly heterogeneous when it comes to forms, two language families stand out because they have very similar markers for certain Evidential categories: Quechuan (discussed below) and Nadahup. The two Nadahup languages in the sample have the same marker for SECONDHAND: the Hup clitic =*mah* and the Dâw particle *māh*. Interestingly, Puinave, an unclassified language, also has a SECONDHAND suffix -*ma*.

- (6.25) Hup (Nadahup; Epps 2008, 654)

tih ham-tég=maḥ

3SG go-FUT=SCND

'He'll go (he or another said so).'

- (6.26) Dâw (Nadahup; Martins 2004, 488)

māh tūg tēh

SCND howler.monkey TOP

'They say that it was the howler monkey which responded. (orig. 'Dizem que foi o guariba quem respondeu.')

- (6.27) Puinave (unclassified; Girón 2008, 376)

ja-dik-ma mowejin

3SG-have-SCND dangerous.spirit

'It is said that there is a devil.' (orig. 'Dizque hay diablo.')

The affiliation of Puinave is still unresolved, but it has been suggested that it originated as Nadahup and considerably changed under Arawakan influence (cf. Girón 2008). According to Muysken (2012), Puinave and Nadahup (and Arawakan) speakers are in contact. If the theory about the Nadahup origin of Puinave is correct, *-ma* is one piece of evidence for this, or Puinave has borrowed *-ma* through contact (in which case it is still possible that it is Nadahup). There is no marker corresponding in both form and function found in the Arawakan languages in the sample, although all of them do have SECONDHAND marking.

The relationship between the Nambikwaran SECONDHAND markers and the SECONDHAND particle *hetau* in Urarina is not clear. Although *hetau* is formally remarkably similar to *-setaw* and *-satau*, according to Olawsky (2006, 503) it is a lexicalization of the reassurance marker *=tau* and the SECONDHAND clitic *=he*. The fact that Urarina and the Nambikwaran languages are not geographically close also speaks against diffusion of forms. Nothing is known about a possible internal division of *-setaw* or *-satau*, although it is possible that at least *-setaw* could be divided into the SECONDHAND suffix *-se?* and an unknown form. The following paragraphs discuss Evidentiality in selected language families in detail: Arawakan (6.5.2), Tupían (6.5.3), Cariban (6.5.4), and Quechuan (6.5.5).

6.5.2 Arawakan

The Arawakan language family varies widely regarding Evidentiality marking: it ranges from marking none at all (Paresi), to marking SECONDHAND only (Yanesha', Baure, Apurinã), to Tariana, which is one of the two languages in the sample marking NON- FIRSTHAND, SECONDHAND, VISUAL, INFERRED, and ASSUMED.

There could have been a direct borrowing of form between Arawakan and Chococoan: Tariana and Apurinã have SECONDHAND markers that are similar to those of Embera and Epena Pedee, both Chococoan languages: *=pida*, *-pira*, *-pida*, and *pída*, respectively. A very

tentative conclusion based on the scarce evidence available points toward an Arawakan origin which will now be discussed.

Tariana has three markers for SECONDHAND fused with PRESENT, REMOTE PAST, and RECENT PAST: the clitics =*pida*, =*pidana*, and =*pidaka*, respectively (Aikhenvald 2003, 302f.). The PRESENT SECONDHAND clitic is also used in the SECONDHAND IMPERATIVE which is a calque from Tucanoan (ibid. 304). Aikhenvald argues that the “specific inferred” Evidential is also a calque from Tucanoan and that this is due to areal induced change (ibid. 287); in fact, the complete Tense/ Evidential system in Tariana is a result of areal diffusion from East-Tucanoan. The form =*pida* existed at first as a general SECONDHAND marker and only acquired a PRESENT specification after contact (ibid. 322). The Tucanoan languages in the sample, Desano and Cubeo, indeed have SECONDHAND markers, but neither of them have a =*pida* form. However, two other unrelated languages with that form are Embera (Chocoan) and Epena Pedee (Chocoan). Embera has a SECONDHAND suffix -*pida* for verbal reports, but also features the suffix -*mana* for narratives, especially legends (Mortensen 1999, 86f.). Epena Pedee uses the particle *pída* to indicate reported knowledge (Harms 1994, 177).⁵ The SECONDHAND marker in Apurinã (Arawakan) is the suffix -*pira* which is apparently not obligatory (Facundes 2000, 317), but the other Arawakan languages in the sample do not have a similar SECONDHAND marker or none at all (a quick search of Chamí (variation of Embera) in Aguirre Licht (1999) yielded no results for either Evidentiality, -*pida*, or -*mana*).

(6.28) Embera (Chocoan; Mortensen 1999, 86)

o-s^hi-pida
make-PST-RPRT
‘[He told me that] he made one.’

(6.29) Epena Pedee (Chocoan; Harms 1994, 177)

josé-pa táama pee-t^haa-hí pída pak^hurú-pa
José-ERG snake kill-OBJ-PST SCND tree-INST
He said that José killed a snake with a club.’

(6.30) Apurinã (Arawakan; Facundes 2000, 317)

ata kema-pira-ta-ru
1PL listen-SCND-VBLZ-3M.OBJ
‘We heard rumours.’

The fact that four languages of two families have almost the same form for the same meaning decreases the possibility of individual development. Additionally, Tariana and Chocoan are reasonably close geographically, though not in the immediate vicinity. Adelaar

⁵ No other SECONDHAND form could be found in Harms (1994), but there is a homophonous form *pída* ‘even’. The exact relationship to the SECONDHAND marker is unknown.

& Muysken (2004, 50) mention that “some surviving groups (e.g. the Chocoan Emberá, the Cuna, the Paéz) have been remarkably expansive in recent times”. The fact that the SECONDHAND markers are formally exactly the same points toward a relatively recent borrowing, or they would be expected to have undergone language-internal change. Because Embera at this moment has two SECONDHAND markers, it is likely that *-mana* is the older form which has been pushed into a more restricted role of legend marking when *-pida* entered the grammar. This is at the moment the only clear factor in favor of an Arawakan origin of **pida*.

Table 6.6: Evidentiality marking in Arawakan

Language/ category	NFRST	SCND	VIS	INFR	ASM
Paresi	-	-	-	-	-
Yanesha	-	=o' (auxiliary <i>ahuo'</i> ?)	-	-	-
Baure	-	=hi	-	-	-
Apurinã	-	=pira	-	-	-
Tariana	=mha, =mahka, =mhana	=pida, =pidaka, =pidana	=naka, =ka, =na	=nihka, =nhina	=sika, =sina

6.5.3 Tupían

According to Rodrigues (1999, 119), Evidentiality is not a major characteristic in Tupían, but, nevertheless, it is found in all Tupían languages in the sample. Especially the SECONDHAND markers suggest common origins: Jensen (1998, 553) lists several cognates in Tupían that have Evidential or Epistemic Modal meanings: the Wayampi verification Modals *ipo* and *jê* also occur in Tupinambá (*ipo* ‘maybe’), Tembê (*ipo* ‘uncertainty or doubt’), Guarayu (*aipo* ‘hearsay’); and in Tembê (*je* ‘they say’), Kamaiurá (*je* ‘they say’) and Urubú-Kaapor (*je* ‘hearsay’). The Guajará hearsay particle *ze* and possibly also the Karo SECONDHAND marker *tã* are cognate. Rodrigues & Cabral (2012, 527) suggest a *dicendi/faciendi* verb (**k²e* ‘to say’) as source for these forms. Additionally, Suruí (also Tupían) has a hearsay particle *ĩyã* (Rodrigues 1999, 119), which may be related to the Cocama-Cocamilla SECONDHAND clitics =ya/ =a and =ia/ =a. Rodrigues & Cabral (2012, 561) state that although Tupían languages differ greatly with regard to Modality, every language has at least one Modal and that this expresses “hearsay”. This is confirmed in the present study.

Three of the five languages in the sample that mark VISUAL are Tupían: Karo, Sateré-Mawé, and Kamaiurá. It is unknown whether their markers are related, but it is possible that VISUAL markers are a Tupían characteristic.

Table 6.7 Evidential marking in Tupían⁶

Language	Family	FRST	NFRST	SCND	QUOT	VIS	INFR	ASM
Emérillon	Tupían, Guaraní	-lako	-	-	-	-	-	-
Tapiete	Tupían, Guaraní	=mba	=nda	-	-	-	-	-
Cocama- Cocamilla	Tupían, Guaraní	-	-	=ya/ =a, =ía/ =a	ay, na	-	-	-
Kamaiurá	Tupían, Guaraní	-	-	je	-	ehe-he	-	-
Karitiana	Tupían	-	-	-sarit	-	-	-	-
Sateré-Mawé	Tupían	-	-	so	-	neke	-	-
Mekens	Tupían	-	-	-	ke, kaat	-	-	-
Karo	Tupían	-	-	tə	-	topə	aket	igā, memā
Suruí	Tupían, Monde	?	?	īyā	?	?	?	?
Guajajára	Tupían, Guaraní	?	?	ze	?	?	?	?
Guarayu	Tupían, Guaraní	?	?	aipo	?	?	?	?
Tembé	Tupían, Guaraní	?	?	je	?	?	?	?
Urubú- Kaapor	Tupían, Guaraní	?	?	je	?	?	?	?

6.5.4 Cariban

The Cariban languages in the sample stand out because two of them do not feature Evidentiality at all (Tiriyó and Panare) and the third one only has SECONDHAND (Hixkaryana). Evidentiality seems to play a minor role in Cariban. Cariban languages have a large number of particles which partially have been classified as “verification (i.e. evidential)” particles by Derbyshire (1999, 53) (see also Derbyshire 1985, 127). Hoff (1986) presents an analysis of 17 particles he claims are Evidentials in Cariban, based on a Cariban speaker from Surinam. His definition of Evidentiality, however, differs somewhat from more contemporary definitions of Evidentiality. According to Hoff (ibid. 49-50), Evidentials “indicate what kind of evidence is available for the reliability of the statement in which they are used”, referring to Givón (1982). Sources of evidence in Hoff’s study are either speaker-internal or the outside world and either strong or weak; the speaker is emotionally involved and makes a statement about his assessment of the truth of the proposition. His particles do not correspond to the notion of Evidentiality adopted here. For example, for Hoff (ibid. 51-53) the prefix *kī-* indicates ‘strong introspective’ commitment of the speaker, i.e. the speaker does not have

⁶ Additional data for Guajajára, Guarayu, Tembé, and Urubú-Kaapor from Jensen (1998); for Suruí from Rodrigues (1999).

any evidence but is sure about what he says. With our present knowledge of Evidentials, this would rather be analyzed as Epistemic Modality rather than Evidentiality.

In that light, it is remarkable that Hixkaryana does have a marker from the domain of Evidentiality, and not Modality: the SECONDHAND particle *tí* (its absence can indicate direct witness) (Derbyshire 1985, 127). Additionally, if a speaker uses this particle s/he rejects any responsibility of the proposition. The question is whether Hixkaryana developed this particle internally, i.e. if we can infer the grammaticalization path, and whether this was induced by contact with another SECONDHAND marking language, or if this form was even completely borrowed. Because SECONDHAND markers logically often develop from verbs denoting ‘say’⁷ or similar meanings, this may be the origin for *tí* as well. However, the Hixkaryana verb ‘say’ is *ka* (ibid. 23). Another possibility is that it developed from a former Modal, but, as Aikhenvald (2004, 151) states: “Such particles [modals] do not usually give rise to grammatical evidentials”. For now, I can only conclude that the Hixkaryana SECONDHAND particle did not grammaticalize from the verb ‘say’ *ka* (although there may another verb ‘say’), and that it presents an anomaly in the otherwise devoid of Evidentiality Cariban languages in this sample.

6.5.5 Quechuan

Quechuan is very homogeneous in its Evidentiality marking with one apparent exception, the INFERRED/ ASSUMED marker in Cuzco Quechua (which may turn out not to be an exception at all). In order to widen the scope of the original sample, additional data from Adelaar with Muysken (2004) and Hintz (2007) is added. Before starting the discussion, a terminological note is in order. Adelaar with Muysken (2004) give several examples of markers they call both ‘evidential’ and ‘validational’, but which only in parts concur with the definitions of Evidential categories presented above. The markers they discuss have been taken into consideration for the present discussion only insofar as they can be described as evidential according to the definition used in this chapter.

Almost all languages in table 6.8 below have a FIRSTHAND suffix *-mi*, and a SECONDHAND OR NONFIRSTHAND marker *-shi*. According to Cole (1985, 164), *-shi* as a hearsay marker occurs only in non-Ecuadorian Quechuan. The marker *-shi* in Imbabura Quechua has also undergone semantic shift from hearsay to speculation (ibid. 165), but not for all speakers; a shift that is not evident in the other Quechuan languages.

The cumulative INFERRED/ ASSUMED suffix *-chá* in Cuzco Quechua merits attention, because it has Epistemic Modal elements. It occurs in the same verbal slot as the other Evidentials (173), but, additionally to coding that a certain amount of reasoning is involved, it also expresses a value of uncertainty: “*-chá* is not purely evidential, indicating that the speaker arrives at his or her statement by reasoning, but also encodes that the speaker is less than 100% certain that the proposition is true” (Faller 2002, 177). However, Faller (2002) rejects

⁷ Cf. the SECONDHAND suffix *-ti* in Tsafiki which derives from the verb *ti* ‘say’ (Dickinson 2002, 103).

an analysis of *-chá* as pure Epistemic Modal. Some other Quechuan languages feature similar markers, often with Epistemic functions. In Imbabura Quechua (Cole 1985) there is a DUBITATIVE marker *-chá(ri)*, but also an interrogative *-chá*. Cole (1985) does not mention any evidential uses of *-chá*.

(6.31) Imbabura Quechua (Quechuan; Cole 1982, 155, 170)

- (a) *shamu-ngui-man-chá*
 come-2-conditional-inter
 ‘Would you come?’
- (b) *kaya-pash-chari shamu-nga*
 tomorrow-even-DUB come-FUT
 ‘Perhaps tomorrow he’ll come.’

Weber (1996, 425) presents an Evidential suffix *-chi* in Huallaga Quechua, which is not treated here as Evidential, because no convincing evidence could be found for this. That suffix does encode conjecture on the part of the speaker, but also shifting the responsibility away from the speaker, as well as the speaker’s subjective assessment of the proposition. It is well possible that it is similar to *-chá* in Cuzco Quechua. Just as Faller (2002, 188) claims that *-chá* is both an Evidential and Epistemic Modal, *-chi* in Huallaga Quechua could also be both. In South Conchucos Quechua, *-chi* denotes both Epistemic (validational) and Evidential (conjectural) values (Hintz 2007, 70), although its primary function is that of validation: “[t]he conjectural *-chi* is used primarily to indicate the speaker’s attitude that the statement s/he is making is probably true (a validational function). It can also be considered to have an evidential function, in which the speaker’s *assumption* is the source of the evidence” (ibid. 69, Hintz’s emphasis). Formally similar Epistemic markers occur in Pacaraos, Tarma, Ayacucho, and Salasaca Quechua (see also chapter 5.10.2 *Quechuan*), but a more detailed study is necessary to investigate the relationship of these with Evidentiality.

An approach to this problem has been suggested by Hintz (2012), who presents an additional notion necessary for the classification of the Evidential markers in South Conchucos Quechua: mutual vs. individual knowledge. Whereas *-shi* remains a simple reportative, the pragmatic and semantic functions of the other markers can then be better understood. Besides *-shi*, South Conchucos Quechua has four more Evidential markers: *-mi*, *-cha*, *-chi*, and *-cher*. Hintz (ibid.) argues that these can best be described by discourse principles, i.e. the participants encode whether they believe the knowledge is shared by the other participants. In that light, *-mi* and *-cha*: both refer to direct witness, but the former encodes “assertion of individual knowledge” and the latter “assertion of mutual knowledge”. Similarly, *-chi* and *-cher* are indirect witness markers, but the first indicates “individual conjecture” and the second “appeal to mutual knowledge”. A similar analysis is given by Hintz (2012) for Evidential markers in Sihuas Quechua. It would be very interesting to see whether shared vs. mutual knowledge can explain anomalies in other Quechuan and non-Quechuan languages as well.

Table 6.8: Evidentiality in Quechuan⁸

	FRST	NFRST	SCND	QUOT	VIS	INFR	ASM
Imbabura Quechua	<i>-mi, -mari</i>	-	<i>-shi (-nin)</i>	-	-	<i>(-chá(ri)?)</i>	-
Huallaga Quechua	<i>-mi</i>	<i>-shi</i>	-	-	-	<i>(-chí?)</i>	-
Cuzco Quechua	<i>-mi</i>	-	<i>-si</i>	-	-	<i>-chá</i>	<i>-chá</i>
Tarma Quechua	<i>-m</i>	?	<i>-š</i>	?	?	<i>(-č?)⁹</i>	?
Ayacucho Quechua	<i>-mi</i>	?	?	?	?	<i>(-cha¹⁰?)</i>	?

6.6 STABILITY OF EVIDENTIALITY

This section evaluates the stability of Evidentiality and draws upon the studies by Wichman & Holman (2009) and Wichmann & Kamholz (2008) (see section 3.10 for an introduction to stability of typological features and the stability of Tense, 4.11 for Aspect, 5.13 for Modality, and 7.5 for a final discussion). The Evidential features included in Wichmann & Holman (2009) include “Semantic Distinction of Evidentiality” and “Coding of Evidentiality”. The former is ranked unstable and the latter very unstable. In the present data, Evidentiality has been shown to diffuse, for instance to Tariana from Tucanoan, but also to be changing within one language family, as in Quechuan. Although Quechuan probably has a shallow time-depth, the Evidential systems have already diverged to a certain extent. Like Aspect and Modality, and to a certain extent also Tense, the present data confirm that Evidential categories are not stable and prone to change internal to language families as well as to diffusion beyond family boundaries. According to Wichman & Holman (*ibid*, 33), categories that are conditioned by pragmatic factors are generally unstable, which may also be an explanation for the instability of Evidentiality.

6.7 SUMMARY

This chapter investigated the typological, geographical, and genealogical patterns of Evidentiality marking in the sample. The languages were analyzed according to FIRSTHAND, NON-FIRSTHAND, SECONDHAND, VISUAL, QUOTATIVE, INFERRED, and ASSUMED Evidentiality. Definitions were given in section 6.2, which are closely modeled on Aikhenvald (2004).

⁸ Data for Tarma and Ayacucho Quechua from Adelaar & Muysken (2004); for Ayacucho Quechua additionally from Soto Ruiz (1976).

⁹ This suffix is given as ‘conjecture’ but also as ‘dubitative’ (Adelaar & Muysken 2004, 210). It has been added here because of its similarity to Cuzco Quechua INFERRED and ASSUMED, but its Evidential status is not confirmed.

¹⁰ According to Soto Ruiz (1976, 124), *-cha* in Ayacucho Quechua means conjecture and possibility.

A relatively high number of languages in the sample mark Evidentiality morpho-syntactically: 49 out of 63 exhibit at least one of the Evidential categories chosen for this study. Of these 49 languages, 22 mark a direct witness (i.e. *FIRSTHAND* and/ or *VISUAL*), and 46 an indirect witness Evidential (i.e. *NON-FIRSTHAND*, and/ or *SECONDHAND*, and/ or *QUOTATIVE*, and/ or *INFERRED* and/ or *ASSUMED*). An intersecting set of 20 languages mark both direct and indirect witness. The most frequently marked category is *SECONDHAND* (38), followed by *FIRSTHAND* (17), *INFERRED* (15), *NON-FIRSTHAND* (11), *ASSUMED* (10), *VISUAL* (5), and *QUOTATIVE* (2). 23 languages in the sample mark one Evidential category, 11 mark 2, 9 mark 3, 4 mark 4, and 2 mark 5. One language, Emérillon, marks direct Evidentiality only. According to De Haan (2011) and Torres Sánchez (2008), there are very few languages world-wide only marking direct Evidentiality.

The languages in the sample can be grouped into Evidential systems of marking one Evidential category to five categories. Systems with one category prominently include *SECONDHAND* and are the most frequent. There are two languages marking five categories: Tariana and Mamaindê.

The geographical distribution of morpho-syntactic Evidentiality is to a high degree heterogeneous. Evidentiality marking occurs everywhere on the continent with no apparent clusters. There is no particular patterning of Evidentiality in the Guaporé-Mamoré, as has been claimed, but because the sample is too small this cannot be stated with certainty.

Evidentiality is in certain cases genealogically conditioned. Cognates are found in Quechuan (*FIRSTHAND* *-mi*, *SECONDHAND* *-shi*, and probably *INFERRED* *-cha/ -chi*), Tupían (*SECONDHAND*), and Nadahup (*SECONDHAND*). Cariban and Arawakan, on the other hand, are completely heterogeneous.

7. DISCUSSION AND CONCLUSIONS

7.1 INTRODUCTION

The geographical distribution of morpho-syntactic TAME marking in the sample is very heterogeneous, as was shown in chapters 3-6. There are few clear regional patterns. The genealogical distribution fares slightly better. Especially the TAME systems of Quechuan and Guaycuruan form distinct units. In contrast, families with more time-depth like Tupían and Arawakan are very heterogeneous. The overall lack of clear patterns may be conditioned by the temporal instability of TAME marking as I will argue in section 7.5. Whereas the geographical and genealogical distributions of the individual TAME categories investigated in this study were discussed in chapters 3 to 6, this chapter attempts to shift from individual categories towards a comprehensive perspective.

For that purpose, it was computed how many TAME categories are marked by each language. In total, a maximum of 35 categories could possibly be marked by a single language (five for Tense, eight for Aspect, 15 for Modality, and seven for Evidentiality). Table 7.2 presents three different values for each supercategory per language. The first column shows the number of marked categories. The second column shows the percentage of marking of the respective category. Column three shows the percentage of marking compared to all categories, but with the supercategories having the same weight. Thus, Tense, Aspect, Modality, and Evidentiality value 25% each. The third to last column presents the sum of all marked categories per language. The second-to-last and last columns present the percentages of how many categories out of 35 are marked, with the difference that the latter equates the categories (each 25%), whereas the former does not. For example, Chimila marks one Tense category, one Aspectual, and zero Modality and Evidentiality (respective first columns). In relation to the respective category, Chimila marks 20% Tense, 25% Aspect, and zero percent Modality and Evidentiality (respective second columns). In relation to all categories Chimila marks 5% Tense, and 6.25% Aspect (respective third column). In total, Chimila marks 11.25% of possible 100% when all supercategories are equated (very last column), and 8.3% when they are not equated (second-to-last column). The languages are ranked from least to most marking of the second-to-last column. For the geographical distribution of the number of TAME markings based on the third-to-last column see map 7.1.

Results show that languages mark between three and 23 categories, or between 8.6% and 65.7% (equated)/ 8.325% and 60.275% (not equated). The ranking differs with regard to whether the supercategories are equated or not, in some cases quite drastically. For example, without equation Panare scores 22.9%, but 35.8% with equation.

Sections 7.2 and 7.3 present an overview of the geographic and genealogical distribution of morpho-syntactic TAME marking in the sample and offer tentative conclusions with regard to regional patterns and the time-depth of families. Section 7.4 presents a statistical analysis of the data putting to the test Bhat's (1999) prominence theory. Section 7.5 resumes

the discussion of the temporal stability of TAME features, and section 7.6 presents future paths of research to complement this study.

7.2 GEOGRAPHICAL DISTRIBUTION IN THE SAMPLE

7.2.1 Overview

What are the geographical patterns of TAME marking in the sample, including the well-known though disputed macro-areas of the Andes and Amazonia and several other linguistic areas: the Içana-Vaupés (Aikhenvald 1999, Epps 2005), the Guaporé-Mamoré (Crevels & Van der Voort 2008), the Upper Xingú (Seki 1999), and the Chaco (Comrie et al. forthcoming)? No clear evidence for an Andean vs. Amazonian typological split could be found (see below), but the Chaco and the Vaupés distinctively mark a very low and very high number of TAME categories, respectively. In agreement with its genealogical diversity the Guaporé-Mamoré shows a broad variety in marking the number of TAME categories. As there are only two languages from the Upper Xingú in the present sample (Trumai and Kamaiurá), it can only be stated here that their tendency is towards the middle range of marking.

In general, the number of TAME markers seems to be highest in the middle range of South America (broadly in Amazonia and adjacent regions), and lower towards the south (the Chaco) and the northwestern corner (the Chibchan sphere). Languages with the middle range of TAME marking especially cluster from central Peru to southern Colombia, whereas the highest number occurs in and around the Vaupés and the Guaporé-Mamoré. That is also where the two major clusters of FRUSTRATIVE OCCUR (see also Aikhenvald 2012 and Overall 2012 for FRUSTRATIVE in Amazonia). Additionally, three of the four languages marking ANTERIOR OCCUR in these two areas (and the fourth in north Peru), as well as the two QUOTATIVES. All other categories do not show such conclusive geographical patterning.

7.2.2 Andes vs. Amazonia

One hypothesis dominant in SA typology in the last decades is the division of SA languages into two different typological profiles that are reflected by the geographical and cultural split of Andean vs. Amazonian. The theory behind this split is the assumption that there is enough evidence for a general, not just linguistic, split in SA that is also reflected in language: “[i]t has been a common practice among linguists working on South American languages to make an intuitive distinction between ‘Amazonian’ and ‘Andean’ languages on the assumption that there would be two different language types corresponding to these labels. Obviously, this distinction is largely fed by geographical and cultural considerations” (Adelaar 2008, 23). Several arguments have been put forth in favor of an Amazonian area (e.g. Dixon & Aikhenvald 1999) and have been rejected (Payne 2001). Constenla (1991) did not find conclusive evidence for an Andean area, and according to Torres Sánchez (2008) there is no significant distribution of Evidentiality along an Andean vs. Amazonian divide.

The analysis in these previous studies is hampered by several factors: there is no consensus of the exact extension of either ‘Andean’ or ‘Amazonian’, and the studies

currently available dealing with this split often focus on a specific group of languages or features, and are not *per se* comparable (cf. Van Gijn, 2013). As of yet, there has been no extended study of all languages in a defined region and of all parts of grammar with respect to an Andean vs. Amazonian split, and so the terms ‘Andean’ and ‘Amazonian’ continue to be used inconsistently, sometimes applied in a geographical sense, sometimes used in order to refer to typologically defined subgroupings. In this light, it seems hardly appropriate to add another study, but because TAME categories have been conspicuously absent from the discussion so far, the following paragraphs take the opportunity to fill the gap.

The majority of linguistic features proposed for an Andean vs. Amazonian profile are not related to TAME categories. The proposed list of features for an Amazonian linguistic area from Dixon & Aikhenvald (1999, 9) includes only one feature relating to TAME, i.e. that the TAME markers are usually suffixes: “[m]ost verbal categories (e.g. tense, aspect, modality, direction) are expressed through optional suffixes”. Adelaar (2008) refers to the rich verbal morphology of Andean languages and mentions that a number of Andean (e.g. Quechuan, Aymaran, Tsafiki, Paez) as well as some pre-Andean languages (Tucanoan and Panoan) feature Evidential systems. However, Evidentiality in this study has been shown to frequently occur throughout the continent without favoring a special region (see section 6.3). It does occur in Andean languages in the sample (e.g. in Quechuan, Aymaran), but also in Amazonia (e.g. Hixkaryana, Desano) and other parts of SA (e.g. in Tapiete in the Chaco).

This lack of TAME features in the Andean vs. Amazonian split discussion (with the exception of Evidentiality) possibly results from the fact that, at least according to the present study, there is no such split. Instead, the data from this sample suggest a broader division for the three categories PRESENT, FRUSTRATIVE, and VISUAL which are almost exclusively absent from western SA. Overtly marked PRESENT occurs almost exclusively in the eastern part of South America; in other words, it is absent along the Andean range. There are two exceptions: Embera and Aymara do have overt PRESENT marking. Aymara encodes PRESENT in NONFUTURE suffixes and Embera in PRESENT suffixes; all of these have varying additional Aspectual and Evidential meanings. The same broad divide occurs in the FRUSTRATIVE data (with the exceptions of Aguaruna and Mapuche) and VISUAL. Of these three categories, the most likely candidate for an actual areal feature is the FRUSTRATIVE, which has already been proposed by Aikhenvald (2012) as a probable feature of Amazonia and/ or the Vaupés (see section 5.10.2). Because VISUAL is very rare in the sample, its absence in western SA is not conclusive.

The present study shows that there is little evidence for an Andean or Amazonian profile in the data for most of the TAME categories, but rather an eastern vs. western SA distinction of morpho-syntactic PRESENT, FRUSTRATIVE, and VISUAL marking. Birchall (2013) presents a reasoned analysis of argument realization patterns across the two macro-areas and argues for a more general east-west split. Krasnoukhova (2012) finds a more fine-grained areal distribution for noun phrase morphosyntax. In general, Amazonia features a higher number of marked TAME categories whereas the Andean region ranges in the

middle. Whether this holds on a more detailed level and whether this is perhaps due to the stability of the categories will have to be shown in the future.

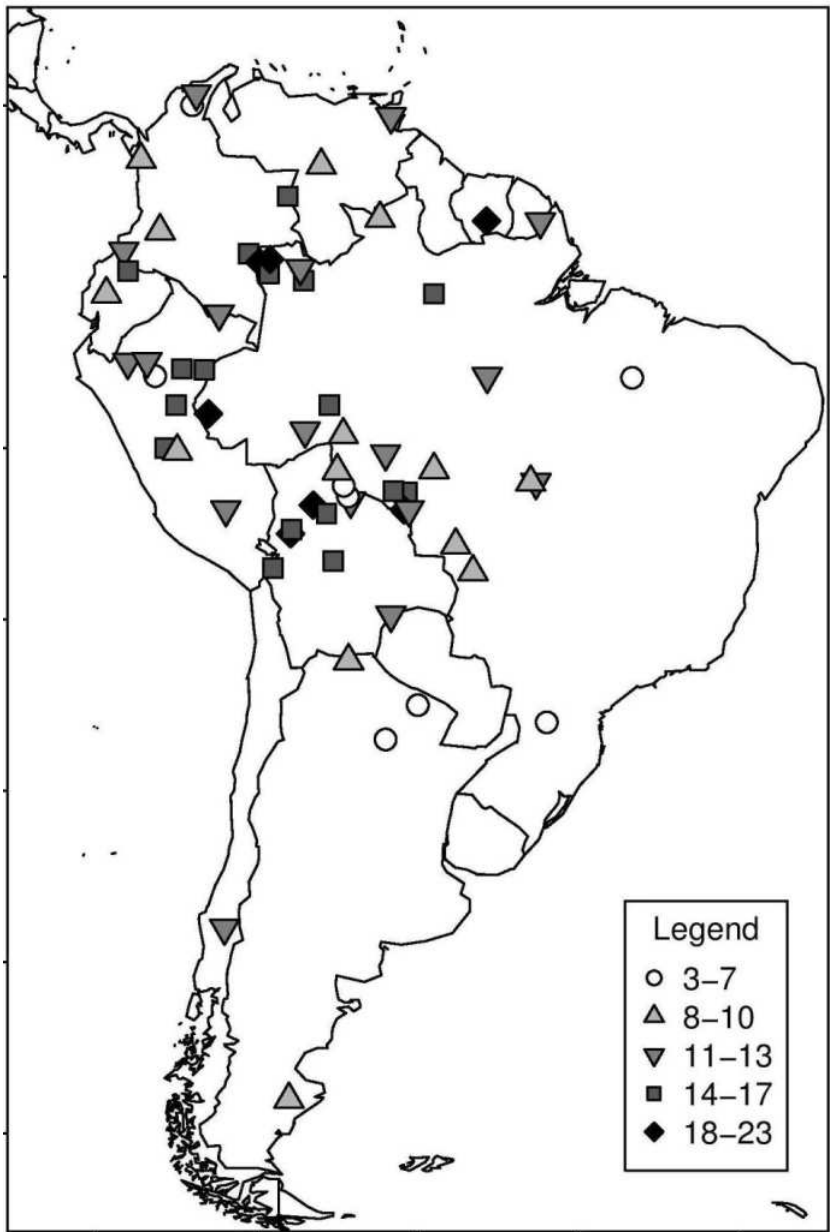
7.2.3 The Chaco

Until recently not enough data was available for a detailed study of the exact relationship between (unrelated) languages in the Chaco (cf. Adelaar with Muysken 2004, 499). Now, Comrie et al. (forthcoming) propose that the Chaco constitutes a linguistic area on the basis of phonological and morpho-syntactic features in Pilagá, Wichí, Tapiete, and Vilela. They furthermore suggest that the Guaycuruan and Matacoan-Mataguayan languages form a core area and Vilela and Tapiete a peripheral area. In the sample for this study, languages in the Chaco are Pilagá and Mocoví (Guaycuruan), Wichí (Matacoan), and Tapiete (Tupían). It was observed in chapters 3 to 6 that the Guaycuruan languages are very poor in morpho-syntactic TAME marking: they do not mark any Tense or COMMAND types, only one Modal (DESIDERATIVE), a few Aspectuals (CONTINUATIVE, HABITUAL, COMPLETIVE), and one Evidential (NON-FIRSTHAND in Pilagá). Guaycuruan is also the family with the least TAME marking in the whole sample. By comparison, Wichí and Tapiete are slightly more elaborate, but still range at the lower end of marking. This can be seen on map 7.1, where Pilagá and Mocoví are in the group with the lowest number of marked TAME categories, Wichí in the second-lowest, and Tapiete in the middle group.

Although it is necessary to include more languages in the sample, it can be said that the languages in the Chaco in the present sample are notable for their low degree of morpho-syntactic TAME marking. Lowest is Guaycuruan, followed by Wichí and Tapiete, which has most marking of the four. Possibly, the core area of the Chaco is characterized by low to almost no TAME marking which increases towards the peripheral area. However, as Krasnoukhova (2013) shows, languages in this area have such deictic systems that often encode TAME reference.

Map 7.1: Number of marked TAME categories

Number of marked TAME categories



7.3 GENEALOGICAL DISTRIBUTION IN THE SAMPLE

This section presents an overview of TAME marking according to language families with regard to their homogeneity (concordance of category marking) and cognacy (concordance of markers of the same category). The following paragraphs briefly summarize the findings for Macro-Gêan, Arawakan, Tupían, Cariban, and Quechuan TAME marking before turning to their possible ranking on a time-depth scale.

In the four Macro-Gêan languages in the sample relatively little homogeneity could be found. There is no conformity in Aspect and Modality marking, but they do seem to share (poor) Tense morphology with usually only FUTURE marking. Two languages mark Evidentiality (SECONDHAND).

The Arawakan languages in the sample show great variety in all TAME categories. They are overall poor in Tense morphology (except for Tariana) but have complex Modal systems. Most of them mark Evidentiality (SECONDHAND). When Tense is marked, it is usually FUTURE. There are some cognates, for example PERFECTIVE *-pe in Baure and Apurinã and IRREALIS *-sia in Baure, Paresi, Apurinã, and Terëna. Tariana differs greatly from the other Arawakan languages, especially in that it developed a complex Evidential system due to contact with Tucanoan.

Tupían also shows a high variety of TAME marking, with a certain number of cognate morphemes. Tense marking ranges from fairly simple to quite elaborate. Aspect, Modality, and Evidentiality systems are likewise heterogeneous. Reduced Evidential systems occur in all Tupían languages and especially the SECONDHAND and Epistemic Modals show a relatively high degree of cognacy. Other cognates can be established for example for *potal in FUTURE markers in Tapiete and Emérillon, *-pab for COMPLETIVE in all Guaraní languages except Tapiete, FRUSTRATIVE in almost all Tupían languages, and *e- (sg), *pe- (pl) for IMPERATIVE in Tapiete, Tupinamba, and Emérillon (all Guaraní). Three of the five VISUAL markers in the sample occur in Tupían.

In this sample, Cariban languages are characterized by their high degree of fusion between Tense, Aspect, and Modality, but not Evidentiality. This may explain their relative homogeneity. For example, all Cariban languages mark remoteness degrees in the PAST but not in the FUTURE, and all prominently mark PERFECTIVE and IMPERFECTIVE. Evidentiality seems to play a minor role in Cariban: only Hixkaryana has an Evidential marker (the SECONDHAND particle *tí*). Cognates occur for DESIDERATIVE and PURPOSIVE, and as reflexes of NONPAST *-ya in IMPERFECTIVE (Panare), PAST PERFECTIVE and PAST IMPERFECTIVE (Hixkaryana), and in FUTURE PERFECTIVE, PRESENT IMPERFECTIVE, and PAST IMPERFECTIVE (Tiriyó).

All three Quechuan languages in the sample are highly homogeneous in TAME marking. For example, there is only minor variation in Tense marking (in the FUTURE), almost all languages have a FIRSTHAND suffix *-mi*, a SECONDHAND/ NONFIRSTHAND marker *-shi*, and an IRREALIS suffix *-man*. In comparison, Aspect marking seems more varied. Although all languages mark PERFECTIVE, they differ widely in which other Aspectual categories are marked, and the degree of cognacy is lower than in the other categories. Except for some Aspectuals, almost all markers are apparently cognate.

To summarize, Arawakan and Tupían are both highly heterogeneous with regard to which categories are marked, but Tupían has a higher degree of cognacy. Cariban TAME marking is more homogeneous than Arawakan or Tupían but less so than Quechuan, which shows the highest degree of homogeneity in the sample. Macro-Gêan does show some homogeneity in Tense but not in the other categories. This perhaps reflects different time-depths of these families, according to which Quechuan would have the shallowest time-depth, Tupían and Arawakan the deepest followed by Macro-Gêan, and Cariban with a medium one. The other language families in the sample rank, with regard to concordance of category marking and cognacy, between the two extremes presented by Quechuan (shallow) and Arawakan/ Tupían (deep). For example, according to Loos (1999, 227), Panoan has “a fairly shallow time-depth and recent expansion and split”. The three Panoan languages in the sample (Matses, Shipibo-Konibo, and Yaminahua) do show some concordance in Tense marking, but are otherwise not as homogeneous as e.g. Quechuan. The Guaycuruan languages are probably nearest to the Quechuan degree of homogeneity, although their striking characteristic is the absence of most TAME marking rather than a high degree of cognacy and homogeneity. Ranging towards the medium level, like Cariban, is Nambikwaran.

Considering the percentage of TAME marking (see table 7.1), it is also observable that most of the language families show a wide range, with some internal subregularities in Tupían and Arawakan. The most homogeneous family is again Quechuan. All three Quechuan languages mark around 40% of TAME both equated and non-equated. The Arawakan languages (without Tariana) differ between 25% and 37% (non-equated) and 21% and 30% (equated). Tariana marks 66% and 58%, respectively, which is quite outside the range of the other Arawakan languages. Tariana is also the language with the highest percentages of TAME marking in the complete sample. The Tupían range of marking is even more varied than the Arawakan one: Karitiana ranks lowest with 23% /27%, and Cocama-Cocamilla highest with 45% and 48%. Karo also has a high equated value (48%), but shares the non-equated value of 34% with Sateré-Mawé, Kamaiurá, and Tapiete. This is due to its high number of Evidentiality marking (4). Interestingly, there is no grouping of Guaraní within Tupían according to the number of marked TAME categories; Guaraní is just as diverse as non-Guaraní.

The ranking of the language families in the sample presents a tentative order from shallow to deep time-depth according to homogeneity and cognacy of TAME marking. More research is necessary to validate this order, by e.g. expanding the sample and comparing the results with other domains of grammar. It would be interesting to see whether members of Arawakan, Cariban, Macro-Gêan, Quechuan, and Tupían not in the sample follow the percentage ranges presented above. It would be expected that Quechuan languages show values around 40%, but that values for the other families differ just as in the sample.

Table 7.1: Sum of TAME categories marked by language families

Language	Family.	TOTAL % (OF 35)	TOTAL % OF ALL CATEGORIES EQUATED
Yanesha'	Arawakan	25.7	27.975
Paresi	Arawakan	28.6	27.675
Apurinã	Arawakan	34.3	21.11
Baure	Arawakan	37.1	30.825
Tariana	Arawakan	65.7	58.225
Panare	Cariban	22.9	35.8
Hixkaryana	Cariban	42.9	44.575
Tiriyó	Cariban	51.4	49.1
Kaingang	Macro-Gêan	17.1	19.575
Timbira	Macro-Gêan	17.1	17.675
Bororo	Macro-Gêan	22.9	22.975
Rikbaktsa	Macro-Gêan	28.6	30.975
Cuzco Quechua	Quechuan	37.1	40.175
Huallaga Quechua	Quechuan	40	41.025
Imbabura Quechua	Quechuan	42.9	41.2
Karitiana	Tupían	22.9	26.725
Nheengatú	Tupían, Guaraní	31.4	29.4
Sateré-Mawé	Tupían	34.3	32.9
Kamaiurá	Tupían, Guaraní	34.3	29.95
Tapiete	Tupían, Guaraní	34.3	39.95
Karo	Tupían	34.3	48.34
Emérillon	Tupían, Guaraní	37.1	34.575
Mekens	Tupían	40	39.575
Cocama-Cocamilla	Tupían, Guaraní	45.1	48.075

Table 7.2: Number of TAME categories marked by each language in the sample

Language	Family	TENSE (MAX. 5)	TENSE % (OF POSSIBLE TENSES)	TENSE % (OF 100% =ALL CATEGORIES EQUATED)	ASPECT (MAX. 8)	ASPECT % (OF POSSIBLE ASPECTUALS)	ASPECT % (OF 100% =ALL CATEGORIES EQUATED)	MODALITY TOTAL (MAX 15)	MODALITY % (OF POSSIBLE MODALS)	MODALITY % (OF 100% =ALL CATEGORIES EQUATED)	EVIDENTIALITY (MAX. 7)	EVIDENTIALITY % (OF POSSIBLE EVIDENTIALS)	EVIDENTIALITY % (OF 100% =ALL CATEGORIES EQUATED)	NUMBER OF MARKED CATEGORIES (MAX. 35)	TOTAL % (OF 35)	TOTAL % OF ALL CATEGORIES EQUATED
Chimila	Chibchan	1	20	5	2	25	6,25	0	0	0	0	0	0	3	8,6	11,25
Mocoví	Guaycuruan	0	0	0	1	12,5	3,125	1	6,6	1,6	1	14,3	3,6	3	8,6	8,325
Pilagá	Guaycuruan	0	0	0	3	37,5	9,375	1	6,6	1,6	0	0	0	4	11,4	10,975
Kanoé	Unclassified	2	40	10	1	12,5	3,125	2	6,6	3,3	0	0	0	5	14,3	16,425
Kaingang	Macro-Géan	1	20	5	3	37,5	9,375	1	6,6	1,6	1	14,3	3,6	6	17,1	19,575
Timbra	Macro-Géan	1	20	5	3	37,5	9,375	2	13,3	3,3	0	0	0	6	17,1	17,675
Itomama	Unclassified	1	20	5	2	25	6,25	2	13,3	3,3	1	14,3	3,6	6	17,1	18,15
Munichi	Unclassified	1	20	5	3	37,5	9,375	3	13,3	3,3	0	0	0	7	20	19,375
Yanam	Yanomaman	0	0	0	2	25	6,25	4	26,6	6,6	2	28,6	7,1	8	22,9	19,95
Kartiana	Tupían	3	60	15	1	12,5	3,125	3	20	5	1	14,3	3,6	8	22,9	26,725
Panare	Cariban	4	80	20	4	50	12,5	2	13,3	3,3	0	0	0	8	22,9	35,8
Bororo	Macro-Géan	1	20	5	3	37,5	9,375	3	20	5	1	14,3	3,6	8	22,9	22,975
Wari'	Chapacuran	4	80	20	2	25	6,25	2	13,3	3,3	0	0	0	8	22,9	29,55
Yanesha'	Arawakan	2	40	10	3	37,5	9,375	3	20	5	1	14,3	3,6	9	25,7	27,975
Embera	Chocoran	5	100	25	3	37,5	9,375	2	13,3	3,3	1	14,3	3,6	9	25,7	41,275
Tehuelche	Chonan	3	60	15	2	25	6,25	2	13,3	3,3	0	0	0	9	25,7	24,55

Wichi (Mataco)	Matacoan	3	60	15	1	12.5	3.125	5	33.3	8.3	0	0	0	0	9	25.7	26.425
Trumai	Unclassified	2	40	10	3	37.5	9.375	4	26.6	6.6	1	14.3	3.6	10	28.6	29.575	
Tsafiki	Barbacoan	0	0	0	2	25	6.25	6	40	10	2	28.6	7.1	10	28.6	23.35	
Paresi	Arawakan	2	40	10	3	37.5	9.375	5	33.3	8.3	0	0	0	10	28.6	27.675	
Rikbaktsa	Macro-Géan	3	60	15	3	37.5	9.375	4	26.6	6.6	0	0	0	10	28.6	30.975	
Nasa Yuwe	Paezan	0	0	0	4	50	12.5	5	33.3	8.3	1	14.3	3.6	10	28.6	24.4	
Awa Pit	Barbacoan	2	40	10	2	25	6.25	7	46.6	11.6	0	0	0	11	31.4	27.85	
Nheengatú	Tupían, Guaraní	1	20	5	4	50	12.5	5	33.3	8.3	1	14.3	3.6	11	31.4	29.4	
Ika	Chibchan, Aruak	2	40	10	2	25	6.25	5	33.3	8.3	2	28.6	7.1	11	31.4	31.65	
Sabanê	Nambikwaran	3	60	15	2	25	6.25	3	20	5	3	42.9	10.7	11	31.4	36.95	
Urarina	Unclassified	0	0	0	3	37.5	9.375	6	40	10	2	28.6	7.1	11	31.4	26.475	
Apuriná	Arawakan	1	20	5	4	50	12.5	6	40	10	1	14.3	3.6	12	34.3	21.11	
Sateré-Mawé	Tupían	1	20	5	4	50	12.5	5	33.3	8.3	2	28.6	7.1	12	34.3	32.9	
Kamaurá	Tupían, Guaraní	1	20	5	2	25	6.25	7	46.6	11.6	2	28.6	7.1	12	34.3	29.95	
Tapiete	Tupían, Guaraní	4	80	20	2	25	6.25	4	26.6	6.6	2	28.6	7.1	12	34.3	39.95	
Karo	Tupían	3	60	15	1	12.5	3.125	4	26.6	6.6	4	57.1	14.2	12	34.3	48.3	
Mapuche	Araucanian	1	20	5	2	25	6.25	8	53.3	13.3	1	14.3	3.6	12	34.3	28.15	
Miraña	Boran	4	80	20	3	37.5	9.375	4	26.6	6.6	2	28.6	7.1	13	37.1	43.075	
Baure	Arawakan	0	0	0	5	62.5	15.625	7	46.6	11.6	1	14.3	3.6	13	37.1	30.825	
Eméřillon	Tupían, Guaraní	2	40	10	3	37.5	9.375	7	46.6	11.6	1	14.3	3.6	13	37.1	34.575	
Cuzco Quechua	Quechuan	2	40	10	3	37.5	9.375	4	26.6	6.6	4	57.1	14.2	13	37.1	40.175	
Aguaruna	Jivaroan	3	60	15	2	25	6.25	8	53.3	13.3	0	0	0	13	37.1	34.55	
Warao	Unclassified	3	60	15	3	37.5	9.375	6	40	10	1	14.3	3.6	13	37.1	37.975	
Jarawara	Arawan	3	60	15	2	25	6.25	6	40	10	3	42.9	10.7	14	40	41.95	

Huallaga	Quechuan	2	40	10	5	62.5	15,625	5	33.3	8.3	2	28.6	7.1	14	40	41,025
Quechua																
Mosetén	Mosetenan	2	40	10	3	37.5	9,375	6	40	10	3	42.9	10.7	14	40	40,075
Dâw	Makuan	3	60	15	5	62.5	15,625	5	33.3	8.3	1	14.3	3.6	14	40	42,525
Mekens	Tupian	3	60	15	3	37.5	9,375	7	46.6	11.6	1	14.3	3.6	14	40	39,575
Movima	Unclassified	3	60	15	1	12.5	3,125	9	60	15	1	14.3	3.6	14	40	36,725
Hikaryana	Cariban	4	80	20	3	37.5	9,375	7	46.6	11.6	1	14.3	3.6	15	42.9	44,575
Imbabura	Quechuan	2	40	10	4	50	12.5	7	46.6	11.6	2	28.6	7.1	15	42.9	41.2
Quechua																
Puinave	Unclassified	3	60	15	3	37.5	9,375	8	53.3	13.3	1	14.3	3.6	15	42.9	41,275
Matses	Panoan	4	80	20	2	25	6,25	7	46.6	11.6	3	42.9	10.7	16	45.1	53.3
Shipibo-Konibo	Panoan	4	80	20	4	50	12.5	4	26.6	6.6	4	57.1	14.2	16	45.1	53.3
Cocama-Cocamilla	Tupian-Guaraní	4	80	20	3	37.5	9,375	7	46.6	11.6	2	28.6	7.1	16	45.1	48,075
Kwaza	Unclassified	3	60	15	3	37.5	9,375	9	60	15	1	14.3	3.6	16	45.1	42,975
Aymara	Aymaran	4	80	20	4	50	12.5	6	40	10	3	42.9	10.7	17	48.6	53.2
Hup	Makuan	4	80	20	5	62.5	15,625	5	33.3	8.3	3	42.9	10.7	17	48.6	54,625
Cubeo	Tucanoan	5	100	25	3	37.5	9,375	5	33.3	8.3	4	57.1	14.2	17	48.6	56,875
Yurakaré	Unclassified	1	20	5	4	50	12.5	9	60	15	3	42.9	10.7	17	48.6	43.2
Tiriyó	Cariban	4	80	20	4	50	12.5	10	66.6	16.6	0	0	0	18	51.4	49.1
Yamnahua	Panoan	4	80	20	4	50	12.5	9	60	15	1	14.3	3.6	18	51.4	51.1
Desano	Tucanoan	4	80	20	4	50	12.5	7	46.6	11.6	3	42.9	10.7	18	51.4	54.8
Leko	Unclassified	4	80	20	5	62.5	15,625	7	46.6	11.6	3	42.9	10.7	19	54.3	57,925
Mamaindé	Nambikwaran	4	80	20	3	37.5	9,375	8	53.3	13.3	5	71.4	17.6	20	57.1	60,275
Caviteña	Tacanan	3	60	15	5	62.5	15,625	11	73.3	18.3	1	14.3	3.6	20	57.1	52,535
Tariana	Arawakan	4	80	10	5	62.5	15,625	9	60	15	5	71.4	17.6	23	65.7	58,225

7.4 TAME PROMINENCE

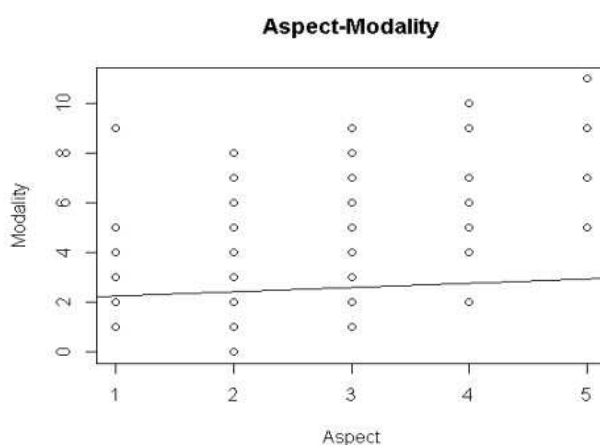
Throughout this study, it was shown that TAME categories are related on a semantic level, and that often a marker of one category (e.g. PERFECTIVE) has secondary functions of another category (e.g. PAST). The same tendency in a sample of languages in India has led Bhat (1999) to believe that some languages tend to give prominence to either Tense, Aspect, or Mood and that the prominent category has functions of the other:

“languages that give greater prominence to one of these categories appeared to view concepts belonging to the other two categories in terms of their prominent category. For example, mood-prominent languages appeared to view temporal and aspectual notions in terms of the modal category, whereas aspect-prominent and tense-prominent languages appeared to view modal (and other) notions in terms of the category of aspect and tense respectively” (ibid. 7).

A first impression is that the prominence theory also holds for the present sample, although Bhat’s study is not restricted to morpho-syntactic marking and does not include Evidentiality. The following discussion investigates whether this holds up to statistical analysis as well. It evaluates the correlation co-efficients of TAME categories and whether they can be shown to be statistically significant. The null hypothesis constitutes that, because TAME categories are inherently related, they should show a high correlation. But can one category (Tense, Aspect, Modality, Evidentiality) predict another one? Is there a ranking of which category better predicts one or the other? The following analysis calculates the correlation co-efficients (r-value) and significance (p-values) levels of Tense, Aspect, Modality, and Evidentiality. First, it was calculated how many subcategories are marked per language per category (see table 7.2, explained in section 7.1). Based on these numbers, the correlation co-efficient was calculated pair-wise for the whole sample, and a significance test carried out to establish whether the co-efficient is likely due to random variation (i.e. not significant). The following nine correlations were tested: Tense with Aspect, Tense with Modality, Tense with Evidentiality, Aspect with Modality, Aspect with Evidentiality, and Modality with Evidentiality, but also combinations: Tense/ Aspect combined with Modality/ Evidentiality combined, and likewise Tense/ Modality with Aspect/ Evidentiality, and Aspect/ Modality with Tense/ Evidentiality. The results for the respective first columns of TAME in table 7.2 are shown in table 7.2, ranked by r-value from highest to lowest. The correlations of the first four pairs are quite probably not due to random variation, but the lower five pairs do not show significant r-values, i.e. their correlations are probably more due to chance than any other factor. For example, it is apparent that some factor other than chance is responsible for the correlation between the number of marked categories of Aspect and Modality, but there is no apparent correlation between Aspect and Tense. The correlations show that there is a trend that when one category goes up, the other (of the pair) also goes up. For instance, when the number of morpho-syntactic Aspect marking increases, the number of Modality marking also increases

(see figure 7.2). In the first four pairs the p-value is significant, i.e. in these cases the correlation is not due to chance. In the latter five pairs, the p-value is not significant, i.e. there is no correlation between the pairs of the kind that when one increases, the other also increases.

Figure 7.1: Linear regression of the number of Aspect and Modality marking



But what we would expect from the prominence theory in any case would be a relationship of the kind 'when one category increases, the other decreases'. For example, in a language with a high degree of Tense marking, Aspect marking should have a rather low degree in comparison. Nevertheless, no such relationship for the five correlations with or without significant p-values could be established. All linear regression models of these are inconclusive. The same calculations were carried out for the respective second and third TAME columns of table 7.2, i.e. whether there are correlations between the values as percentages, and between the values when all four categories are equated. The correlations for the second columns expectedly do not vary much from those of the first column. The results from the third column are mostly not significant, with the exceptions of the correlations of Aspect/ Modality ($r=0,3382669$) and Modality/ Evidentiality ($r=0,2628028$). All of the significant or not significant correlations regression models are inconclusive.

There are some languages in the sample that do appear to show prominence differences, though. For example, Embera marks about 42% of all possible TAME categories (when each supercategory=25%), of which Tense marks 25%, in comparison to Aspect (9.3%), Modality (3.3%), and Evidentiality (3.6%). Thus, Embera seems to be a Tense prominent language. But there are also languages with two or more categories that are almost equal, e.g. Trumai marks both Tense and Aspect with about 10%, and Tariana marks Aspect, Modality, and Evidentiality with between 15% and 18%. 'Prominence' in these cases is not a fixed value, but rather a subjective evaluation.

As a result, there is no statistical signal for a correlation of the number of morpho-syntactic TAME categories according to prominence theory, but there are several significant correlations of the type that when one category increases, the other also increases. This is unexpected and it has yet to be shown what could explain this pattern. Without resorting to calculations, however, there do seem to be languages with a more prominently marked category, although this evaluation is subjective. A further step now would be to calculate the co-efficients of the different TAME subcategories (e.g. PERFECTIVE with PAST) to see whether the prominence theory holds on a lower level than Tense, Aspect, Modality, and Evidentiality, and to carry out the same study on non-morpho-syntactic marking in the same sample.

Table 7.2: Correlations of TAME marking

Correlation pairs	r-value	p-value
Tense/ Aspect - Modality/ Evidentiality	0,4440736	0,002666 (p<0,01)
Tense/ Modality - Aspect/ Evidentiality	0,380859	0,002074 (p<0,01)
Aspect - Modality	0,3635983	0,003399 (p<0,01)
Aspect/ Modality - Tense/ Evidentiality	0,2772911	0,02779 (p<0,05)
Tense - Modality	0,228363	0,07184 (p>0,05)
Tense - Evidentiality	0,2069224	0,1037 (p>0,05)
Modality - Evidentiality	0,1672296	0,1902 (p>0,05)
Aspect - Evidentiality	0,1424128	0,2655 (p>0,05)
Tense - Aspect	0,1390638	0,277 (p>0,05)

7.5 STABILITY

It has been suggested that the relatively high heterogeneity of TAME marking in the sample is due to its temporal instability, which in turn is evoked by its susceptibility to internal change as well as diffusion. In sections 3.10, 4.11, 5.13, and 6.6 the results of the study by Wichmann & Holman (2009) regarding stability of individual TAME features were compared to the present study. This section now seeks to give an overview of the stability of TAME marking in general.

Table 7.3 shows the assessment of the phylogenetic stability of TAME features by Wichmann & Holman (2009) in a sample of 2488 languages world-wide. They calculated the probability that a given language remains unchanged with respect to the feature during 1000 years on the basis of WALS features. The results in table 7.3 represent the calculation on the full set of WALS features. The percentages in the fourth column refer to the amount of similarity between two languages of the same family, i.e. 100% means that the feature is identical in related languages but not in unrelated ones, and 0% means that the feature is just as similar between unrelated languages as in related ones. The stability ranking in column three refers to the four way classification very stable (100%-75%)-stable (75%-50%)-

unstable (50%-25%)-very unstable (25%-100%), with the highest ranking feature of the whole study = 100%.¹

The general tendency of TAME features to remain unchanged in two related languages within 1000 years is low. Most of the features rank from unstable to very unstable, i.e. there is no observable stability of features within one family that would be different from unrelated languages. As a conclusion, the distribution of TAME features is not, or only partially, genealogically conditioned, but other factors such as diffusion and internal change due to e.g. grammaticalization must play an important role.

Interestingly, Tense has different stability values. PAST Tense ranks very stable, but FUTURE is unstable, i.e. PAST Tense is a better indicator of genealogical relationships than FUTURE. In the present study, however, the morpho-syntactic marking of FUTURE is a better indicator of genealogical relationships than PAST. For example, the two language families Tupían and Arawakan are more homogeneous in marking FUTURE than PAST: nine out of nine Tupían languages mark FUTURE, but only six mark PAST; and four out of five Arawakan languages mark FUTURE, but only two mark PAST. Although this points towards FUTURE being more stable than PAST in related SAILS (contrary to Wichmann & Holman 2009), one still needs to take into account the fact that FUTURE is marked very often in this sample. It is just as well marked in related and non-related languages, i.e. FUTURE marking occurs similarly in unrelated languages and related ones which equals a low stability which is in line with Wichmann & Holman.

There is a connection between temporal instability of TAME as presented by Wichmann & Holman (2009) and the heterogeneity of TAME marking both genealogically and geographically in this study. I suggest that the origins are the close inherent semantic relationships of TAME categories which are, for example, evident in the high predisposition towards grammaticalization. Future research hopefully adds TAME categories not (yet) considered by WALS to the list of features, which are expected to be mostly unstable, but Modality and Evidentiality more so than Tense and Aspect.

¹ The highest ranking feature in the study is “sex-based and non-sex-based gender systems”(80.8%), the lowest is “obligatory possessive inflection” (-24.9%) (Wichmann & Holman 2009, 44).

Table 7.3: Temporal stability of TAME features according to Wichmann & Holman (2009)

Feature	WALS feature number	Stability	Stability (%)
The Past Tense	66	Very stable	52.4
Purpose Clauses	125	Stable	48.3
Reduplication	27	Stable	36.2
Perfective/ Imperfective Aspect	65	Stable	36
Situational Possibility	74	Unstable	30.3
Semantic Distinctions of Evidentiality	77	Unstable	28.7
Epistemic Possibility	75	Unstable	28.5
The Future Tense	67	Unstable	26.9
The Morphological Imperative	70	Unstable	26.1
The Prohibitive	71	Unstable	24.2
The Perfect	68	Unstable	22.6
Coding of Evidentiality	78	Very unstable	21.4
Imperative-Hortative Systems	72	Very unstable	18.8
Overlap between Situational and Epistemic Modal Marking	76	Very unstable	7.9

7.6 FUTURE RESEARCH

As mentioned before, this study is but a first step towards a global TAME profile of SAILS. An obvious addition to this study would be the investigation of non-morpho-syntactic TAME expressions, e.g. periphrastic expressions, adverbs, or serial verbs, and whether their distributions would show the same patterns.

There are two dimensions which would profit from expanding the sample: area-specific and family-specific. For instance, one could focus on the linguistic areas of the Guaporé-Mamoré, Içana-Vaupés, or the Chaco, or the macro-areas of the Andes and Amazonia, which have tentatively been discussed in section 7.2. Or zoom in on language families to (i) create samples qualified for testing temporal stability of features, (ii) substantiate family-specific TAME systems, and (iii) compare these to TAME systems of unclassified languages. There are several fields of investigation that have been marginally touched upon here but which deserve more detailed attention. For instance, the unclassified languages in the sample deserve further attention with regard to possible genealogical and geographic relationships to other languages in the sample. Furthermore, there are some promising fields of investigation, for example nominal Tense. Several languages of the SAILS sample show a behavior of temporal markers that has been noted recently also for other languages: Tense appears to be marked not (only) on the verb, but also on other clausal constituents, with scope over an NP and/or the whole clause. The marking of Tense on nouns is known as ‘nominal Tense’ and has been discussed in a cross-linguistic perspective by Nordlinger & Sadler (2004), who have been criticized by Tonhauser (2008). While this phenomenon by itself is already interesting from a typological angle, there may be an areal side to it. Muysken (2008, 83) believes in an “areal effect in the distribution of the Nominal TAM

markers, at least in Amazonia". A closer investigation of the sample according to nominal Tense would show whether this tentative areal distribution holds true.

Finally, it would be interesting to compare this study to archaeological, historical, socio-cultural, anthropological, and genetic accounts. For example, a similar study was carried out concerning the expansion of the Arawak people by Eriksen (2011). His results could be compared against this study, but one could also extend it to other families with available historical sources, such as Quechuan or Tupían.

The topics presented above are but a small part of possible future investigations in the realm of TAME marking in SAILS, which certainly offers a tremendous range of valuable contributions not only to South American studies, but to typology in general.

REFERENCES

- Adelaar, Willem F.H. with the collaboration of Pieter C. Muysken. 2004. *The languages of the Andes*. Cambridge: Cambridge University Press.
- Adelaar, Willem F.H. 2008. Towards a typological profile of the Andean languages. In Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof, eds. *Evidence and counter-evidence. Essays in honour of Frederik Kortlandt. Volume 2: General linguistics*. Amsterdam: Rodopi. 23-33.
- Adelaar Willem F.H. 2012. Historical Overview: Descriptive and comparative research on South American Indigenous Languages. In Campbell & Grondona, eds. 1-58.
- Aikhenvald, Alexandra Y. 1999. Areal diffusion and language contact in the Içana-Vaupés basin, North-West Amazonia. In Dixon & Aikhenvald, eds. 385-416.
- Aikhenvald, Alexandra Y. 2003. *A grammar of Tariana*. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2010. *Imperatives and Commands*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2012. *The languages of the Amazon*. Oxford: Oxford University Press.
- Aguirre Licht, Daniel. 1999. *Embera*. Munich: Lincom Europa.
- Antunes de Araujo, Gabriel. 2004. *A Grammar of Sabanê. A Nambikwaran Language*. LOT Dissertation Series 94. Utrecht: LOT Publications. (Ph.D. dissertation, Vrije Universiteit Amsterdam).
- Alves, Flávia De Castro. 2004. O timbira falado pelos Canela Apaniekrá: uma contribuição aos estudos da morfossintaxe de uma língua Jê. Ph.D. dissertation, Universidade Estadual de Campinas.
- Bacelar, Laércio Nora. 2004. Gramática da língua Kanoê. Ph.D. dissertation, Katholieke Universiteit Nijmegen.
- Bhat, D. N.S. 1999. *The Prominence of Tense, Aspect, and Mood*. Amsterdam/ Philadelphia: John Benjamins.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional Morphology. In Timothy Shopen, ed. *Language Typology and Syntactic Description Vol III*, second edition. Cambridge: Cambridge University Press. 169-240.
- Birchall, Joshua. 2013. Verbal argument marking patterns in South American Languages. In O'Connor & Muysken.
- Bohnenmeyer, Jürgen. 2009. Temporal Anaphora in a Tenseless Language: The Case of Yucatec. In Wolfgang Klein & Ping Li, eds. *The Expression of Time*. Berlin: Mouton de Gruyter. 83-128.
- Brandão, Ana Paula. 2010. Verb morphology in Paresi-Haliti (Arawak). M.A. thesis, University of Texas at Austin.
- Brandão, Ana Paula. 2011. Handout. Tense and Aspect in Paresi-Haliti (Arawak). LinC Colloquium Nijmegen, 29.03.2011.

- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam/ Philadelphia: John Benjamins.
- Bybee, Joan L. & William Pagliuca. 1987. The evolution of future meaning. In Anna Giacalone Ramat, Onofrio Carruba & Giuliano Bernini, eds. *Papers from the 7th International Conference on Historical Linguistics*. Amsterdam/ Philadelphia: John Benjamins. 109–123.
- Bybee, Joan L., Revere Dale Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Bybee, Joan L. 1994. The Grammaticization of Zero. In William Pagliuca, ed. *Perspectives on Grammaticalization*. Amsterdam/ Philadelphia: John. 235–254.
- Campbell, Lyle & Verónica Grondona, eds. *The Indigenous Languages of South America. A Comprehensive Guide*. The World of Linguistics 2. Berlin/ Boston: De Gruyter Mouton.
- Casasanto, Daniel & Lera Boroditsky. 2008. Time in the mind: Using space to think about time. *Cognition* 106: 579–593.
- Cavalcante, Marita Pôrto. 1987. Fonologia e morfologia da língua kaingãng: o dialeto de São Paulo comparado com o do Paranã. Ph.D. dissertation, University of Campinas.
- Cerrón-Palomino, Rodolfo & Juan Carvajal Carvajal. 2009. Aimara. In Crevels & Muysken, eds. 169–214.
- Chung, Sandra & Alan Timberlake. 1985. Tense, Aspect, and Mood. In Timothy Shopen, ed. *Language Typology and Syntactic Description. Grammatical categories and the lexicon*. Cambridge: Cambridge University Press. 202–258.
- Cole, Peter. 1982. *Imbabura Quechua*. Lingua Descriptive Studies. Amsterdam: North-Holland Publishing Company.
- Cole, Peter. 1985. *Imbabura Quechua*. Croom Helm Descriptive Grammars. Croom Helm.
- Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.
- Comrie, Bernard, Lucía A. Golluscio, Hebe González & Alejandra Vidal. Forthcoming. El Chaco como área lingüística.
- Constenla Umaña, Adolfo. 1991. *Las lenguas del área intermedia: Introducción a su estudio areal*. San José: Editorial de la Universidad de Costa Rica.
- Corbett, Greville. 2007. Canonical Typology, Suppletion, and Possible Words. *Language* 83(1): 8–42.
- Crevels, Mily. 2006. Itonama. In Rowicka & Carlin, eds. 159–170.
- Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. In Pieter Muysken, ed. *From linguistic areas to areal linguistics*. Amsterdam/ Philadelphia: John Benjamins. 151–179.
- Crevels, Mily & Pieter Muysken, eds. 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural editores.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Crowell, Thomas Harris. 1979. A grammar of Bororo. Ph.D. dissertation, Cornell University.

- Curnow, Timothy. 1997. A grammar of Awa Pit (Cuaiquer): An indigenous language of south-western Colombia. Ph.D. dissertation, The Australian National University.
- Da Cruz, Aline. 2011. *Fonologia e Gramática do Nheengatú. A língua geral falada pelos povos Baré, Warekena e Baniwa*. LOT Dissertation Series 280. Utrecht: LOT Publications. (Ph.D. dissertation, Vrije Universiteit Amsterdam).
- Dahl, Östen. 1984. Temporal distance: remoteness distinctions in tense-aspect systems. In Brian Butterworth, Bernard Comrie & Östen Dahl, eds. *Explanations for Language Universals*. Berlin: Mouton. 105-122.
- Dahl, Östen & Viveka Velupillai. 2011a. Perfective/ Imperfective Aspect. In Dryer & Haspelmath, eds. Chapter 65. Available online at <http://wals.info/feature/65>. Accessed on 2012-16-08.
- Dahl, Östen & Viveka Velupillai. 2011b. The Past Tense. In Dryer & Haspelmath, eds. Chapter 66. Available online at <http://wals.info/feature/66>. Accessed on 2012-16-08.
- Dahl, Östen & Viveka Velupillai. 2011c. The Perfect. In Dryer & Haspelmath, eds. Chapter 68. Available online at <http://wals.info/feature/65A>. Accessed on 2012-16-08.
- Danielsen, Swintha. 2007. *Baure. An Arawak language of Bolivia (Indigenous Languages of Latin America 6)*. Leiden: CNWS Publications. (Ph.D. dissertation, Radboud Universiteit Nijmegen).
- De Haan, Ferdinand. 2005. Encoding speaker perspective: Evidentials. In Frajzyngier & Rood, eds. *Linguistic diversity and language theories*. Amsterdam/ Philadelphia: John Benjamins. 379-398.
- De Haan, Ferdinand. 2011. Semantic Distinctions of Evidentiality. In Dryer & Haspelmath, eds. Chapter 77. Available online at <http://wals.info/chapter/77>. Accessed on 2012-05-02.
- De Haan, Ferdinand. 2012. Irrealis: fact or fiction? *Language Sciences* 34 (2). 107-130.
- Dendale, Patrick & Liliane Tasmowski. 2001. Introduction: Evidentiality and Related Notions. *Journal of Pragmatics* 33: 339-348.
- Derbyshire, Desmond C. 1979. *Hixkaryana*. Amsterdam: North-Holland Publishing Company.
- Derbyshire, Desmond C. 1985. *Hixkaryana and linguistic typology*. Dallas: Summer Institute of Linguistics.
- Derbyshire, Desmond C. 1986. Comparative Survey of Morphology and Syntax in Brazilian Arawakan. In Derbyshire & Pullum, eds. 469-566.
- Derbyshire, Desmond C. 1999. Carib. In Dixon & Aikhenvald, eds. 23-64.
- Derbyshire, Desmond C. & Geoffrey K. Pullum eds. *Handbook of Amazonian Languages 1*. Berlin/ New York: Mouton de Gruyter.
- Dickinson, Connie. 2002. Complex predicates in Tsafiki. Ph.D. dissertation, University of Oregon.
- Dietrich, Wolf. 2006. La categoría del frustrativo en las lenguas tupí-guaraníes. In Wolf Dietrich & Haralambos Symeonidis, eds. *Guaraní y "Mawetí-tupí-Guaraní": estudios históricos y descriptivos sobre una familia linüística de América del Sur*. Berlin: Lit Verlag. 71-85.

- Dixon, R.M.W. 2004. *The Jarawara Language of Southern Amazonia*. Oxford: Oxford University Press.
- Dixon, R.M.W. & Alexandra Y. Aikhenvald. 1999. *Introduction*. In Dixon & Aikhenvald, eds. 1-22.
- Dixon, R.M.W. & Alexandra Y. Aikhenvald eds. 1999. *The Amazonian Languages*. Cambridge: Cambridge University Press.
- Drude, Sebastian. 2008. Tense, Aspect and Mood on Awetfí verb paradigms. In K. David Harrison, David S. Rood & Arienne M. Dwyer, eds. *Lessons from Documented Endangered Languages*. Amsterdam/ Philadelphia: John Benjamins. 67-109.
- Dryer, Matthew S. & Martin Haspelmath, eds. 2011. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. Available online at <http://wals.info/> Accessed on 2012-11-02.
- Duff-Tripp, Martha. 1997. *Gramática del Idioma Yanasha' (Amuesha)*. Serie Lingüística Peruana 43. Lima, Perú: Instituto Lingüístico de Verano.
- Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink & Angela Terrill. 2008. Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language* 84 (4): 710-759.
- Eberhard, David M. 2009. *Mamaïndê grammar. A Northern Nambikwara language and its cultural context*. LOT Dissertation Series 236. Utrecht: LOT Publications. (Ph.D. dissertation, Vrije Universiteit Amsterdam).
- Epps, Patience. 2005. Areal Diffusion and the Development of Evidentiality. *Studies in Language* 29 (3): 617-650.
- Epps, Patience. 2008. *A grammar of Hup*. Mouton Grammar Library 43. Berlin/ New York Mouton de Gruyter.
- Eriksen, Love. 2011. Nature and culture in prehistoric Amazonia. Using G.I.S. to reconstruct ancient ethnogenetic processes from archeology, linguistics, geography, and ethnohistory. Ph.D. dissertation, Lund University.
- Everett, Caleb. 2006. Gestural, Perceptual, and Conceptual Patterns in Karitiana. Ph.D. dissertation, Rice University.
- Everett, Dan & Barbara Kern. 1997. *Wari'. The Pacaas Novos language of Western Brazil*. London: Routledge.
- Facundes, Sidney da Silva. 2000. The language of the Apurinã people of Brazil. Ph.D. dissertation, SUNY Buffalo.
- Faller, Martina T. 2002. Semantics and Pragmatics of Evidentials in Cuzco Quechua. Ph.D. dissertation, Stanford University.
- Faller, Martina. 2003. The Evidential and Validational Licensing Conditions for the Cusco Quechua Enclitic *-mi*. *Belgian Journal of Linguistics* 16: 7-21.
- Faust, Norma & Eugene Loos. 2002. *Gramática del idioma yaminahua*. Lima: Instituto Lingüístico de Verano.
- Fernández Garay, Ana. 1998. *El Tehuelche. Una lengua en vías de extinción*. Valdivia: Universidad Austral de Chile.

- Fleck, David. 2003. A grammar of Matsés. Ph.D. dissertation, Rice University.
- Frank, Paul Stephen. 1985. A Grammar of Ika. Ph.D. dissertation, University of Pennsylvania.
- Frank, Paul. 1990. *Ika syntax*. Dallas: Summer Institute of Linguistics.
- Gabas Jr., Nilson. 1999. A grammar of Karo, Tupí (Brazil). Ph.D. dissertation, University of California, Santa Barbara.
- Galucio, Ana Vilacy. 2001. The Morphosyntax of Mekens (Tupi). Ph.D dissertation, University of Chicago.
- Gibson, Michael Luke. 1996. *El Munichi. Un idioma que se extingue*. Perú: Instituto Lingüístico de Verano.
- Gildea, Spike. 1989. Simple and Relative Clauses in Panare. Master thesis, University of Oregon.
- Gildea, Spike. 1992. Comparative Cariban Morphosyntax: On the Genesis of Ergativity in Independent Clauses. Ph.D. dissertation, University of Oregon.
- Gildea, Spike. 1998. *On Reconstructing Grammar. Comparative Cariban Morphosyntax*. Oxford: Oxford University Press.
- Gildea, Spike. 2012. Linguistic Studies in the Cariban family. In Campbell & Grondona, eds. 441-494.
- Girón, Jesús Mario. 2008. *Una gramática del Wānsöhöt (Puinave)*. LOT Dissertation Series 185. Utrecht: LOT Publications. (Ph.D. dissertation, Vrije Universiteit Amsterdam).
- Givón, T. 2001. *Syntax: An Introduction*. Amsterdam/ Philadelphia: John Benjamins.
- Goodwin Gómez, Gale. 1990. The Shiriana dialect of Yanam (northern Brazil). Ph.D. dissertation, Columbia University.
- González, Hebe Alicia. 2005. A grammar of Tapiete (Tupí-Guaraní). Ph.D. dissertation, University of Pittsburgh.
- Grondona, Verónica. 1998. A grammar of Mocoví. Ph.D. dissertation, University of Pittsburgh.
- Guillaume, Antoine. 2008. *A Grammar of Cavineña*. Mouton Grammar Library 44. Berlin/ New York: Mouton de Gruyter.
- Guirardello, Raquel. 1999. A Reference Grammar of Trumai. Ph.D. dissertation, Rice University.
- Hammarström, Harald. 2009. Sampling and genealogical coverage in the WALS. *Linguistic Typology* 13 (1): 105-119.
- Hardman, Martha. 2001. *Aymara*. Munich: Lincom Europa.
- Harms, Phillip Lee. 1994. *Epena Pedee Syntax*. Studies in the Languages of Colombia 4. Summer Institute of Linguistics and the University of Texas at Arlington.
- Haspelmath, Martin. 2006. Against Markedness (and What to Replace It With). *Journal of Linguistics* 42 (1): 25-70.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86 (3): 663-687.
- Haude, Katharina. 2006. A grammar of Movima. Ph.D. dissertation, Radboud Universiteit Nijmegen.

- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Hintz, Daniel J. 2012. Building Common Ground. The Evidential Category of Mutual Knowledge. Conference talk, The Nature of Evidentiality, June 14, 2012, Leiden.
- Hintz, Diane M. 2007. Past Tense Forms and Their Functions in South Conchucos Quechua: Time, Evidentiality, Discourse Structure, and Affect. Ph.D. dissertation, University of California, Santa Barbara.
- Jensen, Cheryl. 1998. Comparative Tupí-Guaraní Morpho-syntax. In Desmond C. Derbyshire & Geoffrey K. Pullum, eds. *Handbook of Amazonian Languages 4*. Berlin/ New York: Mouton de Gruyter. 487-618.
- Jensen, Cheryl. 1999. Tupí-Guaraní. In Dixon & Aikhenvald eds. 125-164.
- Jung, Ingrid. 2008. *Gramática del páez o nasa yuwe: Descripción de una lengua indígena de Colombia*. Munich: Lincom Europa.
- Kleiber, Georges. 1998. *Prototypensemantik: eine Einführung*. Tübingen: Narr.
- Klein, Wolfgang. 2009. How Time Is Encoded. In Klein & Li, eds. 39-82.
- Klein, Wolfgang & Ping Li. 2009. *How Time Is Encoded. The Expression of Cognitive Categories 3*. Berlin/ New York: Mouton de Gruyter.
- Krasnoukhova, Olga. 2012. *The Noun Phrase in the Languages of South America*. LOT Dissertation Series 301. Utrecht: LOT Publications. (Ph.D. dissertation, Radboud Universiteit Nijmegen).
- Krasnoukhova, Olga. 2013. Demonstratives in the languages of South America: Redrawing the semantic map. In O'Connor & Muysken.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World, Sixteenth edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Loos, Eugene E. 1999. Pano. In Dixon & Aikhenvald, eds. 227-250.
- Martins, Silvana Andrade. 2004. *Fonología e gramática Dâw*. LOT Dissertation Series 98. Utrecht: LOT Publications. (Ph.D. dissertation, Vrije Universiteit Amsterdam).
- Martins, Silvana & Valteir Martins. 1999. Makú. In Dixon & Aikhenvald, eds. 251-267.
- Meira, Sérgio. 1999. A grammar of Tiriyó. Ph.D. dissertation, Rice University.
- Miller, Marion. 1999. *Desano grammar*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Moore, Denny. 1994. A few aspects of comparative Tupí syntax. *Revista Latinoamericana de Estudios Etnolingüísticos 8* (Special issue by Mary Ruth Wise ed.), *Lingüística tupí-guaraní e caribe. Estudios presentados en el 46 Congreso Internacioal de Americanistas New Orleans 1991*. 151-162.
- Morse, Nancy & Michael Maxwell. 1999. *Cubeo grammar*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Mortensen, Charles. 1999. *A reference grammar of the Northern Emberá languages*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Mueller, Neele. 2013. Language Internal and External Factors in the Development of the Desiderative in South American Indigenous Languages. In O'Connor & Muysken.

- Muysken, Pieter. 1981. Creole Tense/mood/aspect Systems: The Unmarked Case? In Pieter Muysken, ed. *Generative Studies on Creole Languages*. Dordrecht: Foris Publications. 181-199.
- Muysken, Pieter. 2008. Nominal Tense. Time for Further Whorfian Adventures? Commentary on Casasanto. In Peter Indefrey & Marianne Gullberg eds. *Time to Speak: Cognitive and Neural Prerequisites for Time in Language*. Chichester: Wiley-Blackwell. 81-88.
- Narrog, Heiko. 2005. On defining modality again. *Language Sciences* 27 (2). 165-192.
- Narrog, Heiko & Bernd Heine, eds. 2011. *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press.
- Nonato, Rafael Bezerra. 2008. Ainore Boe egore: um estudo descritivo da língua bororo e conseqüências para a teoria de caso e concordância. M.A. Thesis, Universidade Estadual de Campinas.
- Nordlinger, Rachel & Lousia Sadler. 2004. Nominal tense in cross-linguistic perspective. *Language* 80 (4): 776-806.
- Nordlinger, Rachel, & Louisa Sadler. 2008. When is a temporal marker not a tense? *Language* 84 (2): 325-331.
- Nordström, Jackie. 2010. *Modality and Subordinators*. Amsterdam/ Philadelphia: John Benjamins.
- O'Connor, Loretta & Pieter Muysken eds. 2013. *The Native Languages of South America: Origins, development, typology*. Cambridge: Cambridge University Press.
- Olawsky, Knut. 2006. *A Grammar of Urarina*. Mouton Grammar Library 37. Berlin/ New York: Mouton de Gruyter.
- Overall, Simon. 2007. A grammar of Aguaruna. Ph.D. dissertation, La Trobe University.
- Overall, Simon. 2012. Frustrative in Amazonian Languages. Unpublished manuscript. La Trobe University.
- Palmer, F. R. 2001. *Mood and Modality*, second edition. Cambridge: Cambridge University Press.
- Payne, David L. 1991. A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions. In Desmond C. Derbyshire & Geoffrey K. Pullum, eds. *Handbook of Amazonian Languages*, 3. Berlin/ New York: Mouton de Gruyter. 355-500.
- Payne, Doris L. 2001. Review of R.M.W. Dixon & Alexandra Aikhenvald, eds. *The Amazonian Languages*. *Language* 77 (3): 594-598.
- Payne, Doris L. & Thomas E. Payne. 1990. Yagua. In Desmond C. Derbyshire & Geoffrey K. Pullum, eds. *Handbook of Amazonian Languages*, 2. Mouton de Gruyter. 249-474.
- Payne, Thomas E. 1990. Transitivity and Ergativity in Panare. In Doris L. Payne, ed. *Amazonian Linguistics: Studies in Lowland South American languages*. Austin: University of Texas Press. 429-456.
- Payne, Thomas E. 1995. Object Incorporation in Panare. *International Journal of American Linguistics* 61 (3): 295-331.
- Payne, Thomas E. & Doris L. Payne. 2013. *A Typological Grammar of Panare*. Leiden: Brill.

- Reichenbach, Hans. 1948. *Elements of Symbolic Logic*. New York: The Macmillan Company.
- Rodrigues, Aryon D. 1999. Tupí. In Dixon & Aikhenvald, eds. 107-124.
- Rodrigues, Aryon Dall'Igna & Ana Suely Arruda Câmara Cabral. 2012. Tupían. In Campbell & Grondona, eds. 495-574.
- Romero-Figueroa, Andrés. 1997. *A Reference Grammar of Warao*. Munich: Lincom Europa.
- Rosch, Eleanor, Caroline B. Mavis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8 (3): 382-439.
- Rose, Françoise. 2003. Morphosyntaxe de l'emérillon, langue amérindienne de Guyane française. Ph.D. dissertation, Université Lumière Lyon 2.
- Rowicka, Grażyna J. & Eithne B. Carlin, eds. *What's in a Verb? Studies in the verbal morphology of the languages of the Americas*. LOT Occasional Series 5. Utrecht: LOT Publications.
- Rubino, Carl. 2005. Reduplicaton: Form, Function and Distribution. In Bernhard Hurch, ed. *Studies on Reduplication*. Berlin/ New York: Mouton de Gruyter. 11-30.
- Rubino, Carl. 2011. Reduplication. In Dryer & Haspelmath, eds. Chapter 27. Available online at <http://wals.info/feature/27>. Accessed on 2012-01-10.
- Sakel, Jeanette. 2004. *A grammar of Mosestén*. Mouton Grammar Library 33. Berlin/ New York: Mouton de Gruyter.
- Silva, Léia de Jesus. 2005. Aspectos da Fonologia e da Morfologia da Língua Rikbaktsa. M.A. thesis, Universidade de Brasília.
- Silva, Léia de Jesus. 2011. Morphosyntaxe du Rikbaktsa (Amazonie brésilienne). Ph.D. dissertation, Université Denis Diderot – Paris 7.
- Silva, Raynice Geraldine Pereira da. 2010. Estudo Morfosintático da Língua Sateré-Mawé. Ph.D. dissertation, Universidade Estadual de Campinas.
- Schmidtke-Bode, Karsten. 2009. *A Typology of Purpose Clauses*. Amsterdam/ Philadelphia: John Benjamins.
- Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. Nijmegen: Max Planck Institute for Psycholinguistics. (Ph.D. dissertation, Radboud Universiteit Nijmegen).
- Seki, Lucy. 1999. The Upper Xingú as an incipient linguistic area. In Dixon & Aikhenvald, eds. 417-430.
- Seki, Lucy. 2000. *Gramática do Kamaiurá: Língua Tupi-Guarani do Alto Xingu*. Campinas: Editora da Universidade Estadual de Campinas.
- Smeets, Ineke. 2008. *A grammar of Mapuche*. Mouton Grammar Library 41. Berlin/ New York: Mouton de Gruyter.
- Smith, Carlota S. 2008. Time with and without tense. In J. Lecarne & J. Guéron eds. *Time and Modality*, 227-250. Dordrecht: Springer.
- Soto Ruiz, Clodoaldo. 1976. *Gramática Quechua. Ayacucho-Chanca*. Lima: Ministerio de Educación/ Instituto de Estudios Peruanos.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Clarendon Press.
- Telles, Stella & Leo Wetzels. 2006. Evidentiality and Epistemic Mood in Lakondê. In Rowicka & Carlin, eds. 235-252.

- Terraza, Jimena. 2009. Gramática del wichí: fonología y morfosintaxis. Ph.D. dissertation, Université du Québec à Montréal.
- Tonhauser, Judith. 2007. Nominal Tense? The Meaning of Guaraní Nominal Temporal Markers. *Language* 83 (4): 831-869.
- Torres Sánchez, Nadiezdha. 2008. La evidencialidad en las lenguas indígenas americanas. Un enfoque areatipológico. M.A. thesis, Escuela Nacional De Antropología e Historia, Mexico.
- Trillos Amaya, María. 1996. Categorías gramaticales del ette tara. Lengua de los chimilas. Bogotá: CCELA, Universidad de los Andes.
- Ultan, Russell. 1978. The Nature of Future Tenses. In Joseph H. Greenberg, ed. *Universals of Human Language* 3. Stanford: Stanford University Press. 83-125.
- Valenzuela, Pilar M. 2003a. Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan. In Alexandra Y. Aikhenvald & R.M.W. Dixon, eds. *Studies in Evidentiality*. Amsterdam/ Philadelphia: John Benjamins. 33-62.
- Valenzuela, Pilar. 2003b. Transitivity in Shipibo-Konibo grammar. Ph.D. dissertation, University of Oregon.
- Vallejos Yopán, Rosa. 2010. A Grammar of Kokama-Kokamilla. Ph.D. dissertation, University of Oregon.
- Van der Auwera, Johan & Andreas Amman. 2011. Epistemic Possibility. In Dryer & Harnad, eds. Chapter 75. Available online at <http://wals.info/chapter/75>. Accessed on 2012-09-03.
- Van de Kerke, Simon. 2009. Leko. In Crevels & Muysken, eds. 287-332.
- Van der Voort, Hein. 2004. *A Grammar of Kwaza*. Berlin: Mouton de Gruyter.
- Van Gijn, Rik. 2006. A grammar of Yurakaré. Ph.D. dissertation, Radboud Universiteit Nijmegen.
- Van Gijn, Rik. 2013. The Andean foothills and adjacent Amazonian fringe. In O'Connor & Muysken.
- Vidal, Alejandra. 2001. Pilagá grammar (Guaykuruan family, Argentina). Ph.D. dissertation, University of Oregon.
- Villafañe, Lucrecia. 2004. *Gramática Yuki: lengua Tupí-Guaraní de Bolivia*. Tucumán: Ediciones del Rectorado, Universidad Nacional de Tucumán. (Ph.D dissertation, Katholieke Universiteit Nijmegen).
- Weber, David J. 1996. *A grammar of Huallaga (Huánuco) Quechua*. Berkeley: University of California Press.
- West, Birdie. 1980. *Gramática Popular Del Tucano*. Bogotá: Instituto Lingüístico de Verano.
- Wichmann, Søren & Eric W. Holman. 2009. *Assessing Temporal Stability for Linguistic Typological Features*. Munich: Lincom Europa.
- Wichmann, Søren & David Kamholz. 2008. A stability metric for typological features. *Sprachtypologie und Universalienforschung* 61 (3): 251-262.
- Wise, Mary Ruth. 1986. Grammatical characteristics of PreAndine Arawakan languages of Peru. In Derbyshire & Pullum, eds. 567-642.

SUMMARY

This study presents an analysis of morpho-syntactic tense, aspect, modality, and evidentiality (TAME) marking in a sample of 63 South American indigenous languages (SAILS) with regard to typological, geographical, and genealogical distributions. It aims to provide an overview over the occurrence of selected TAME features in South America and seeks to uncover possible patterns that add to our understanding of the distribution of SAILS as it is today.

The basis for the analysis is a questionnaire which was designed i) to capture the major TAME features and ii) to take into account several features of particular interest in SA. The following features are investigated: future, past, present, remoteness degrees in future and past, realis, irrealis, forms of command marking (imperative), intentional, potential, certainty/ uncertainty, dubitative, frustrative, purposive, desiderative, perfective, imperfective, anterior, habitual, continuative, iterative, completive, incompletive, firsthand, non-fristhand, secondhand, quotative, visual, inferred, and assumed. Taken into account are overt morpho-syntactic markers, such as affixes, clitics, and particles. A morpheme qualifies as marker when it fulfills certain characteristics which are demonstrated in the respective TAME chapters.

The language sample was designed to balance geographical and genealogical diversity of SAILS with a focus on certain families and regions, taking the availability of good descriptions into account. The sample represents 25 language families in addition to 11 unclassified languages.

Chapters three, four, five, and six discuss in detail tense, aspect, modality, and evidentiality marking as occurring in the sample. The first part of each chapter is dedicated to a short overview of the topic which is followed by the definitions used in this study. Special sections at the end of each chapter present the typological, geographical, and genealogical distribution of the respective category.

Chapter three investigates tense marking in the sample. It is demonstrated that morpho-syntactic tense marking in the SAILS sample can be divided into tensed and tenseless languages. Most of the languages in the sample have morpho-syntactic tense marking, only seven are tenseless. The most frequently marked category is future, which is marked by 54 languages, followed by past (44), remoteness in the past (28), present (16), and remoteness in the future (11).

A major characteristic of morpho-syntactic tense marking in the SAILS is its heterogeneity of forms which may be explained by the relatively low stability of future and past, marking as calculated in a previous study. Tense marking in all categories is varied in all language families (with more than two members) in the sample with the noticeable exception of Quechuan. In contrast, Arawakan, Macro-Gêan and Tupían languages are widely inconsistent in their tense makers.

Chapter four investigates aspect marking in the sample. No language in the sample is without aspect marking, but also no language marks the possible maximum of eight

aspectual categories. The most frequently marked category is continuative (35), followed by iterative (32) and habitual (31), completive (26), perfective (24), imperfective (22), incomplete (14), and anterior (4). Continuative and habitual are morpho-syntactically marked more frequently in the SAILs than in other regions of the world, whereas perfective/ imperfective are marked by about the same portion in SA.

The marking of aspect within language families varies from being mildly to completely heterogeneous. A few genealogical traits can be identified for Arawakan and Cariban. The fact that morpho-syntactic marking of both aspect and tense is quite heterogeneous is explained by the relationship between the two.

Chapter five investigates modality marking in the sample. Results show that imperative is the most frequently marked category (45 languages), followed by desiderative (39), prohibitive (30), dubitative (28) and irrealis (28), purposive (24), frustrative (23) and hortative (23), certainty (20), potential (17), collective command (16), intentional (13), polite command (11) and jussive (11), and realis (3). Only one language has no modality marking (Chimila), but none marks the maximum possible number (15 including command types; 9 excluding command types).

Several regional clusters of modal marking occur in the sample. The Guaporé-Mamoré area features a considerable variation of high genealogical diversity; on the contrary, the Vaupés region is relatively homogeneous in terms of which modal category is marked. Frustrative occurs only east of the Andes in what roughly constitutes Amazonia, and throughout all language families, but without genealogical signal except for Tupían. Tupían prominently features frustratives which may have two different origins and, possibly, contact with Tupían led to the spread of frustratives in Amazonia. Whereas a single language family was probably responsible for the spread of frustrative, the desiderative was developed by several language families, which explains its wider geographical spread than the spread of frustrative.

Epistemic modality occurs particularly frequently in north-west Brazil and south-east Colombia, including the Vaupés, and several minor clusters in Bolivia and north Peru. Commands occur frequently in the sample in all regions and families, with the exception of Guaycuruan and Macro-Gêan. Many command type markers apparently grammaticalized from the same origins as future markers, or from future markers.

Modality marking varies greatly according to family. Quechuan is very homogeneous in Modality and command type marking, with cognates for almost all categories. On the other hand, Macro-Gêan is very heterogeneous and unified only by its lack of command type marking. Guaycuruan has no command type marking either. The other language families rank in between these two extremes, being relatively heterogeneous with a few sets of cognates. Tupían has frustrative cognates; desiderative occurs in cognates in virtually all families; command type cognates in Cariban and Tupían can be connected to established proto-forms.

Modal features are unstable, because of their relationship with tense, which is also unstable, and its susceptibility to pragmatic factors.

Chapter six investigates evidentiality marking in the sample. A relatively high number of languages in the sample mark evidentiality morpho-syntactically: 49 out of 63 exhibit at least one of the evidential categories chosen for this study. Of these 49 languages, 22 mark a direct witness (i.e. firsthand and/ or visual), and 46 an indirect witness evidential (i.e. non-firsthand, and/ or secondhand, and/ or quotative, and/ or inferred and/ or assumed). An intersecting set of 20 languages mark both direct and indirect witness. The most frequently marked category is secondhand (38), followed by firsthand (17), inferred (15), non-firsthand (11), assumed (10), visual (5), and quotative (2). 23 languages in the sample mark 1 evidential category, 11 mark 2, 9 mark 3, 4 mark 4, and 2 mark 5. One single language, Emérillon, marks direct evidentiality only.

The languages in the sample can be grouped into evidential systems of marking one evidential category to five categories. Systems with one category prominently include secondhand and are the most frequent. There are two languages marking five categories: Tariana and Mamaindê.

The geographical distribution of morpho-syntactic evidentiality is to a high degree heterogeneous. Evidentiality marking occurs everywhere on the continent with no apparent clusters. There is no particular patterning of evidentiality in the Guaporé-Mamoré area, as has been claimed.

Evidentiality marking is in certain cases genealogically conditioned. Cognates are found in Quechuan, Tupían, and Makuan. Cariban and Arawakan, on the other hand, are completely heterogeneous.

Chapter seven consolidates the results from the previous chapters and examines their categories in a comprehensive manner, with special focus on geographical and genealogical distributions, TAME prominence, and temporal stability of TAME features.

SAMENVATTING IN HET NEDERLANDS

Deze studie analyseert de morfosyntactische markering van de categorieën tijd, aspect, modaliteit, en evidentialiteit (TAME) in een steekproef van 63 inheemse talen van Zuid-Amerika. Het proefschrift bespreekt de typologische, geografische en genealogische distributie van deze categorieën. De basisdoelstelling van de studie is een overzicht te geven van de verspreiding en patroonvorming van TAME in Zuid-Amerikaanse talen, en om op die manier de huidige distributie van deze talen proberen te begrijpen en verklaren.

De basis voor de analyse is een vragenlijst ontwikkeld met twee doelen. Het eerste doel is de belangrijkste kenmerken in het domain van TAME weer te geven. Het tweede is aandacht te schenken aan bepaalde eigenschappen die vooral relevant zijn in Zuid-Amerikaanse talen. De volgende kenmerken worden besproken:

Tijd	Aspect	Modaliteit	Evidentialiteit
<i>future</i>	<i>perfective</i>	<i>realis</i>	<i>firsthand</i>
<i>past</i>	<i>imperfective anterior</i>	<i>irrealis</i>	<i>non-firsthand</i>
<i>present</i>	<i>habitual continuative</i>	<i>forms of command</i>	<i>secondhand</i>
<i>remoteness degrees in future and past</i>	<i>iterative</i>	<i>marking (imperative)</i>	<i>quotative</i>
	<i>completive</i>	<i>intentional</i>	<i>visual inferred</i>
	<i>incompletive</i>	<i>potential</i>	<i>assumed</i>
		<i>certainty/ uncertainty</i>	
		<i>dubitative</i>	
		<i>frustrative</i>	
		<i>purposive</i>	
		<i>desiderative</i>	

Er wordt gekeken naar concreet gerealiseerde morfosyntactische markeerders, zoals affixen, clitics, en partikels. Een morfeem wordt als markeerder beschouwd als het voldoet aan bepaalde criteria omschreven en geïllustreerd in de TAME hoofdstukken.

De steekproef van talen is samengesteld op basis van een balans tussen geografische en genealogische diversiteit van Zuid-Amerikaanse talen (en tot op zekere hoogte op basis van beschikbare taalbeschrijvingen van hoge kwaliteit), met een focus op bepaalde families en regio's. Naast 11 ongeclassificeerde talen, worden in de steekproef 25 taalfamilies vertegenwoordigd.

Hoofdstukken drie tot en met zes bespreken in detail de markering van tijd, aspect, modaliteit, en evidentialiteit in de talen van de steekproef. Het eerste deel van elk hoofdstuk is gewijd aan de definitie en een kort overzicht van een categorie. Speciale secties aan het eind van elk hoofdstuk bespreken de typologische, geografische, en genealogische distributie van de betreffende categorie in de talen van de steekproef.

Hoofdstuk **drie** behandelt tijdsmarkering. Er wordt aangetoond dat op basis van morfosyntactische tijdsmarkering Zuid-Amerikaanse talen kunnen worden onderverdeeld

in talen met (*tensed*) en talen zonder tijdsmarkering (*tenseless*). Hoewel het merendeel van de talen in de steekproef morfosyntactische tijdsmarkering heeft, hebben zeven talen deze eigenschap niet. De meest frequent gemarkeerde tijds categorie is toekomst (*future*) (in 54 talen), gevolgd door verleden tijd (*past*) (44), *remoteness in the past* (28), *present* (16), en *remoteness in the future* (11).

Eén van de belangrijkste kenmerken van morfosyntactische tijdsmarkering in Zuid-Amerikaanse talen is de heterogeniteit van vormen, die kan worden verklaard aan de hand van de relatief lage stabiliteit van de markering van toekomst en verleden. Tijdsmarkering in alle categorieën is gevarieerd in alle taalfamilies (met meer dan twee leden) in deze steekproef, met de opvallende uitzondering van de Quechua taalfamilie. In contrast hiermee zijn de Arawak, Macro-Gê en Tupí talen zeer inconsistent in de gebruikte tijdsmarkeerders.

Hoofdstuk **vier** onderzoekt markering van aspect. Geen enkele taal in de steekproef is zonder aspectsmarkering. Tegelijkertijd markeert geen taal het mogelijke maximum van acht aspectuele categorieën. De meest frequent gemarkeerde categorie is *continuative* (35), gevolgd door *iterative* (32) en *habitual* (31), *completive* (26), *perfective* (24), *imperfective* (22), *incompletive* (14) en *anterior* (4). *Continuative* en *habitual* worden morfosyntactisch vaker gemarkeerd in Zuid-Amerikaanse talen dan in de andere talen van de wereld, terwijl de frequentie van *perfective/imperfective* vergelijkbaar is met elders. De markering van aspect binnen taalfamilies varieert van licht tot volledig heterogeen. Een paar genealogische kenmerken kunnen worden geïdentificeerd voor Arawak en Carib talen. Het feit dat de morfosyntactische markering van zowel aspect als tijd vrij heterogeen is, valt te verklaren aan de hand van de relatie tussen de twee categorieën.

Hoofdstuk **vijf** bespreekt markering van modaliteit. De resultaten laten zien dat *imperatief* de meest frequent gemarkeerde categorie is (45 talen), gevolgd door *desiderative* (39), *prohibitive* (30), *dubitative* (28) en *irrealis* (28), *purposive* (24), *frustrative* (23) en *hortative* (23), *certainty* (20), *potential* (17), *collective command* (16), *intentional* (13), *polite command* (11) en *jussive* (11), en *realis* (3). Slechts 1 taal heeft geen markering van modaliteit (Chimila), maar geen enkele taal markeert het maximaal mogelijke aantal categorieën (15 inclusief *command* typen, 9 exclusief *command* typen).

Er kunnen enkele regionale clusters van modaliteitsmarkering worden onderkend. De Guaporé-Mamoré regio toont een hoge genealogische diversiteit; de Vaupés regio, darentegen, is relatief homogeen in termen van de modale categorieën die worden gemarkeerd. Markering van *frustrative* doet zich geografisch alleen ten oosten van de Andes voor – in de regio die grofweg Amazonia omvat – en is in alle taalfamilies te vinden, maar zonder een duidelijk genealogisch signaal, behalve voor de Tupí taalfamilie. *Frustrative* wordt vaak in de Tupí familie gemarkeerd en komt mogelijk voort uit twee verschillende bronnen. Eventueel contact met Tupí kan de verspreiding van *frustrative* in het Amazonegebied verklaren. Terwijl de verspreiding van de *frustrative* markering waarschijnlijk te wijten is aan één taalfamilie, is de markering van *desiderative* waarschijnlijk ontwikkeld door verschillende taalfamilies, hetgeen de bredere geografische spreiding dan die van de *frustrative* verklaart.

Het markeren van epistemische modaliteit komt met name veel voor in het noordwesten van Brazilië en zuidoosten van Colombia, inclusief de Vaupés regio, en een aantal kleinere clusters in Bolivia en noord Peru. *Commands* komen vaak voor in in de talen van alle regio's en taalfamilies, met uitzondering van de Guaycuru en Macro-Gê talen. Veel markeerders van *command* typen zijn blijkbaar gegrammaticaliseerd op basis van dezelfde markeerders voor toekomst.

De markering van modaliteit varieert sterk van familie tot familie. De Quechua talen zijn zeer homogeen in het markeren van modaliteit en *command* typen en hebben cognaten voor bijna alle categorieën. Aan de andere kant zijn Macro-Gê talen zeer heterogeen en delen alleen het ontbreken van markering van *command* typen. Guaycuru talen markeren evenmin *command* typen. De andere taalfamilies liggen tussen deze twee uitersten en zijn relatief heterogeen met een paar sets cognaten. Tupí heeft cognaten voor *frustrative*; *desiderative* komt in cognaten voor in vrijwel alle families; cognaten voor *command* typen in Carib en Tupí kunnen worden gerelateerd aan vastgestelde protovormen. Modale kenmerken zijn instabiel door hun relatie met tijd, die ook instabiel is, en hun gevoeligheid voor pragmatische factoren.

Hoofdstuk **zes** gaat over markering van evidentialiteit. Een relatief groot aantal talen in de steekproef markeert evidentialiteit morfosyntactisch: 49 van de 63 vertonen ten minste een van de evidentialiteitscategorieën uit deze studie. Van deze 49 talen, markeren 22 directe evidentie (d.w.z. *firsthand* en/ of *visual*), en 46 indirecte evidentie (d.w.z. *non-firsthand* en/ of *secondhand*, en/ of *quotative*, en/ of *inferred* en/ of *assumed*). Een groep van 20 talen markeert zowel directe als indirecte evidentie. De meest frequent gemarkeerde categorie is *secondhand* (38), gevolgd door *firsthand* (17), *inferred* (15), *non-first hand* (11), *assumed* (10), *visual* (5), en *quotative* (2). 23 talen in de steekproef markeren 1 evidentialiteitscategorie, 11 talen markeren 2, 9 talen markeren 3, 4 talen markeren 4, en 2 talen markeren 5 categorieën. Eén taal, Emerillon, markeert slechts directe evidentialiteit.

De talen in de steekproef kunnen dus worden gegroepeerd op basis van het markeren van een tot vijf evidentiële categorieën. Systemen met één categorie bevatten altijd *secondhand*, en komen het meest voor. Er zijn twee talen die vijf categorieën markeren: Tariana en Mamaindê. De geografische spreiding van de morfosyntactische markering van evidentialiteit is in hoge mate heterogeen. Markering van evidentialiteit komt overal op het continent voor, zonder duidelijke clusters. Er is geen bepaald patroon van evidentialiteitsmarkering te vinden in de Guaporé-Mamoré regio, zoals wordt gesuggereerd in een eerdere studie. De markering van evidentialiteit is in bepaalde gevallen genealogisch bepaald. Cognaten zijn te vinden in de Quechua, Tupí, en Makú talen. Carib en Arawak zijn daarentegen volledig heterogeen.

Hoofdstuk **zeven**, tenslotte, vergelijkt de resultaten uit de voorgaande hoofdstukken in een breder kader. Er wordt speciale aandacht besteed aan de geografische en genealogische distributie, de relatieve nadruk op één van de TAME categorieën, en de stabiliteit van TAME kenmerken.

CURRICULUM VITAE

Neele Müller was born in Leer, Ostfriesland (Germany), May 28th, 1982. After finishing secondary school in 2002, she studied English Linguistic, Celtic Studies, and German Linguistics at the Phillips-Universität Marburg (Germany) from where she graduated in 2009. During her studies she spent two semesters at Millersville University, PA, and one semester at University College, Cork. Her Magister's thesis focused on Welsh English and the language contact situation between Welsh and English. Besides being professionally interested in endangered languages in general, her private focus is on her native language, East-Frisian.

Additionally, Neele is a hobby-writer of Fantasy and an avid reader of Fantasy and Crime novels, and interested in the early civilizations of, e.g. Sumer, Egypt, and Mesoamerica.