Leiden University
Leiden University Centre for Linguistics

Jens Van Gysel

Temporal Predicative Particles in Sanapaná and the Enlhet-Enenlhet Language Family (Paraguay)

A Descriptive and Comparative Study

Thesis submitted in partial fulfilment of the requirements for the degree of
Master of Arts in Linguistics:
Language Diversity of Africa, Asia and Native America

2017

Supervisor Dr. Mily Crevels
Leiden University Centre for Linguistics
Second reader Prof. dr. Maarten Mous
Leiden University Centre for Linguistics

Acknowledgements

Even though this Master's Thesis is relatively limited in size and scope, the help of many people was essential in its creation. Most importantly, since this thesis is based on original fieldwork, I need to express my deep gratitude and friendship for the Sanapaná people of the La Esperanza community. Without their help, time, energy, kindness, and patience, this work would not have been possible. In particular, I wish to thank don Marino Ortega, the community leader, for his hospitality and willingness to accept me into his community and to allow me to live and work amongst the Sanapaná for two months.

My language teachers as well were essential to the creation of this thesis. Valerio Chávez, Regino Echeverría, Silio Recalde, Doto Teytaro, Cecilio Teytaro, Civito Montes, Esteban López, Cristino Benítez, Pisoa Cabrera, Roberto Álvarez, Enrique Loma and Venancio González on a systematic basis volunteered their time and knowledge to help me learn about the Sanapaná language and customs, and many others offered their help and opportunities to practice my Sanapaná skills more sporadically. Adolf Kehler and his family, John and Patty Stucky, and Luís Martínez and his family also helped make my stay in Paraguay as comfortable and pleasant as possible. To all of them, I extend my sincerest gratitude.

Next, I would not even have made it to La Esperanza if it were not for Carlos Giesbrecht of the NGO Pro Comunidades Indígenas, who agreed to let me do my internship with his organisation; for Manolo Romero, who brought me, my suitcase and my tent from Filadelfia to La Esperanza on his motorbike over 80 kilometers of Chaco roads; and for Hannes Kalisch of Nengvaanemkeskama Nempayvaam En-lhet, who brought me in touch both with Mr. Giesbrecht and with don Marino Ortega, gave me valuable advice before my travel to Paraguay, and helped me grasp the basics of the Enlhet-Enenlhet languages by sharing with me his years-long experience with them. Rodrigo Villagra and John Elliot, too, were kind enough to share their data and insights with me. Additionally, without the funds I received from Leiden University's Sustainable Humanities Internship Fund, the fieldwork that underlies this thesis would not have been possible. To all these people as well, I say a big thank you.

The next person I need to thank is the one who set the whole chain of communication between me and the La Esperanza community in motion by bringing me in touch with Hannes Kalisch: dr. Mily Crevels, my supervisor at Leiden University. For the second year in a row, she guided me through the process of writing an MA thesis. Apart from helping me with her comments to make the end product of this investigation better, she made the investigation itself possible by helping me search for an internship position in South America from the summer of 2016 onwards.

Last, but certainly not least, I need to thank my parents for putting up with my ambitions of being a professional student and for providing me with the financial support I needed to complete this year of study in Leiden, and the moral support needed to finish this thesis. More of this last kind of support was provided by Emily and Sophie, who endured both my complaints of frustration when things were going less well, and my outbursts of joy when I discovered new findings. To all the people mentioned on this page, and those I might have forgotten, I extend a warm thank you.

Agradecimientos

Aunque el tamaño de esta tesis de maestría es relativamente limitado, la ayuda de mucha gente ha sido esencial en su creación. Más importante, visto que esta tesis tiene su base en trabajo de campo, tengo que expresar mi gratitud y amistad para la gente Sanapaná de la Comunidad de La Esperanza. Sin su ayuda, tiempo, energía, amabilidad, y paciencia, este trabajo no habría sido posible. Quiero agradecer particularmente a don Marino Ortega, el líder de la comunidad, por su hospitalidad y su voluntad de aceptarme dentro de su comunidad, y de darme permiso para vivir y trabajar allá durante dos meses.

Mis maestros del idioma Sanapaná también jugaron un papel esencial en la creación de esta tesis. Valerio Chávez, Regino Echeverría, Silio Recalde, Doto Teytaro, Cecilio Teytaro, Civito Montes, Esteban López, Cristino Benítez, Pisoa Cabrera, Roberto Álvarez, Enrique Loma y Venancio González me ofrecieron voluntaria y sistemáticamente su tiempo y conocimiento, enseñándome en el idioma y la cultura Sanapaná. Otros me ofrecieron más esporádicamente su ayuda y oportunidades para praticar el idioma. Adolf Kehler y su familia, John y Patty Stucky, y Luís Martínez y su familia me ayudaron a pasar mi estancia en Paraguay lo más agradable posible. A todos extiendo mi muy sincera gratitud.

Luego, ni siquiera hubiera llegado a La Esperanza si no hubiera podido contar con el ayudo de Carlos Giesbrecht de la ONG Pro Comunidades Indígenas, quien me permitió hacer mis prácticas con su ogranización; de Manolo Romero, quien me llevó a mí, a mi maleta, y a mi carpa de Filadelfia a La Esperanza en su moto sobre 80 kilómetros de caminos chaqueños; y de Hannes Kalisch de Nengvaanemkeskama Nempayvaam Enlhet, quien me puso en contacto tanto con el Sr. Giesbrecht como con don Marino Ortega, quien me dio consejas muy útiles antes de mi viaje al Paraguay, y quien me ayudó a comprender las nociones básicas de los idiomas Enlhet-Enenlhet, compartiendo conmigo su extensa experiencia con ellos. Sin la ayuda financiera del Fondo de Prácticas de Ciencias Humanas Sostenibles de la Universidad de Leiden, el trabajo de campo que forma la base de esta tesis tampoco habría sido posible. A todos ellos también digo muchas gracias.

La siguiente persona que necesito agradecer es la que estableció el red de comunicación entre mí y la Comunidad de La Esperanza, poniéndome en contacto con Hannes Kalisch: dr. Mily Crevels, mi profesor en la Universidad de Leiden. Es el segundo año consecutivo que me ha guiado en el proceso de escribir una tesis de maestría. Ella no solamente me ayudó con sus comentarios que mejoraron el producto final de esta investigación, sino también posibilitó la investigación misma, ayudándome desde el verano de 2016 a conseguir una oportunidad para unas prácticas en América del Sur.

Por último, tengo que agradecer a mis padres para soportar mis ambiciones de volverme un estudiante profesional, y para prestarme la ayuda financiera necesaria para completar este año de estudios en Leiden, y el apoyo moral necesario para terminar esta tesis. Este último tipo de ayuda me prestaron también Emily y Sophie, quienes soportaban tanto mis quejas de frustración cuando no progresaban las cosas de manera muy buena como mis explosiones de felicidad cuando descubría nuevas ideas. A todos y todas nombrados aquí, y los que quizás olvidé, les doy unas gracias muy sinceras.

Abstract

This MA thesis provides the first description and discussion of three temporal predicative particles (*ałta*, *łke* and *hata*) found in Sanapaná, an underdocumented and underdescribed Enlhet-Enenlhet language of the Paraguayan Chaco. Furthermore, it compares the form and function of these particles in Sanapaná to that of their cognates in the other Enlhet-Enenlhet languages: Angaité, Enxet, Enlhet, Guaná and Enenlhet.

As such, this thesis has three main objectives. Firstly, it aims to be a contribution to the descriptive literature on an understudied indigenous language of South America, since it is based on primary data gathered during my own field work. Secondly, it hopes to provide insights into the genetic relations between the languages of the Enlhet-Enenlhet language family. Thirdly, this work hopes to provide data for possible areal or typological studies of the phenomena manifested by these particles: tense-aspect-modality systems, and (nonverbal) predication.

It is argued that *ałta*, when it follows a verb, functions as a prehodiernal past marker. When it follows a word from a different word class, it can locate the referent of a noun in the past (but is not a nominal tense marker), and it makes the inherent predicativity of this word explicit, just as Kalisch (2009) argues that these particles do in Enlhet. *łke*, secondly, is argued to be an immediate aspect marker when it follows a verb, and has the same predicative force when following a word from a different class. Furthermore, it can locate the previous mention of a referent or its previous presence in the extralinguistic context in the past. Lastly, *hata* functions as an indefinite future marker when combined with a verb, and also has this aforementioned predicative force.

Based on these (admittedly limited) Sanapaná data and that of its sister languages, it seems that Unruh & Kalisch' (2003) hypothesis of a Western and an Eastern branch of the family is warranted: Guaná, Sanapaná and Enenlhet cluster together, while Enlhet shares fewer features with these languages. Within the former group, Guaná and Sanapaná seem to be most similar.

Table of Contents

List of Abbreviations and Symbols

A	most agentive argument of transitive clause
ABIT	abitive
ADV	adverb
AFF	affirmative
AUX	auxiliary
C	consonant
CAUS	causative
CF	counterfactual
CPL	complexive
CONJ	conjunction
COP	copula
DEF	definite
DEM	demonstrative
DET	determiner
DIM	diminutive
DIR	directional
DIST	distal
DISTR	distributive
E	event time
E1	starting point of event time
E2	end point of event time
EXT	extensive
F	feminine
FACT	factive
FN	field notes
FOC	focaliser
FUT	future
HOD	hodiernal past
IMM	immediate aspect
IND	indicative
INF	infinitive
INT	interrogative
INTNS	intensifier
IMP	imperative
IMPERS	impersonal
INFR	inferential
IPFV	imperfective
IRR	irrealis
LOC	locative
M	masculine
NEG	negation
NOM	nominative (in glosses); nominal (in the text)
NP	noun phrase
NPST	non-past
O	least agentive argument of a transitive clause
PAS	passive
PFV	perfective
PHOD	prehodiernal past
PL	plural
POSS	possessive
POT	potential
PRES	presentative
PRIM	primative
PRO	pronoun
PROG	progressive
PROSP	prospective
PROX	proximate
PS	present stem(?)
PST	past
R	reference time
REAL	realis
REC	recent past
REFL	reflexive
REP	repetitive
RPT	reportative
S	single argument of an intranstitive clause; speech time
SBJV	subjunctive
SEC	secundive
SG	singular
SP	Spanish loan
STX	stem extension
TAM	tense-aspect-mood
TEMP	temporal
V	vowel
V1	Sanapaná verb form in verb-initial clauses
V2	Sanapaná verb form in non-verb initial clauses
VENT	ventive
x:y	y is an item of category x

x.y	category x and y are fused in one morpheme
x_y	English translation needs multiple words
1	first person
2	second person
3	third person
>	precedes; outranks; changes into
=>	acts upon
’	stress lies on the next syllable

List of Tables and Figures

Tables

Figures

Chapter 1 Introduction

In this MA thesis, I present the first in-depth discussion of the form and function of three particles (*ałta*, *łke* and *hata*) in Sanapaná, an Enlhet-Enenlhet language spoken in the Paraguayan Chaco. I focus specifically on their function in terms of temporal reference and predication. The data on which this discussion is based stem mostly from my own fieldwork in the winter of the 2016-2017 academic year. Furthermore, I compare these particles to their cognates in Enlhet Norte, Guaná and Enenlhet, the three languages of the Enlhet-Enenlhet family for which a certain amount of published linguistic data are available. Based on the findings of the comparison of these three particles, I present some implications for the way in which the genetic relations between these languages should be viewed. In this first chapter, I start by introducing the geographical region of the Gran Chaco, where these languages are spoken, after which I provide a brief overview of the sociolinguistic situation of and the previous research done on the Enlhet-Enenlhet languages. I conclude this introduction by laying out the structure and aims of the rest of this thesis.

1.1 The Gran Chaco region

The Gran Chaco is one of the linguistic hotspots of the South American continent. The region spans roughly 600 000 square kilometres of plains across parts of Bolivia, Brazil, Paraguay, and Argentina. It is bordered in the north by the Mato Grosso plateau, in the east and the west by the Paraguay and Paraná rivers and by the Andean foothills, respectively, and in the south by the Salado river basin (Braunstein & Miller 1999: 1-3).

This region was traditionally inhabited by hunter-gatherer societies, who were contacted by Europeans for the first time in the early 16th century, amongst others during punitive expeditions by Spanish military governors such as Alvar Núñez Cabeza de Vaca (Braunstein & Miller 1999: 5-7). The first attempts to settle the region took place around the turn of the 17th century, but met with considerable resistance on the part of the indigenous inhabitants. These initial settlements were soon abandoned again (Braunstein & Miller 1999: 7-8). Nevertheless, from the 18th century onwards, missionary activity in the Chaco saw a significant increase in intensity, first under the auspices of the Jesuit order, and after their expulsion by the Spanish government under those of the Franciscan order (Krebs & Braunstein 2011: 10-11). Under the influence of these missionaries, many of the indigenous groups underwent a steady process of acculturation, documented for, amongst others, the Paraguayan Enlhet (Kalisch 2012) and the Bolivian ‘Weenhayek (formerly known as Mataco, and known as Wichí in Argentina, Alvarsson

2007). As a result, most indigenous groups nowadays live sedentary lives in the mission settlements and have abandoned many of their traditional cultural practices.

The number of indigenous groups and languages present today in the Chaco region varies depending on which estimate is taken into account. Conservative estimates are found in Braunstein & Miller (1999: 1) and Campbell (2013: 275-8), who posit around 20 languages divided over 6 language families, spoken by around 260 000 indigenous people. A more generous estimate is provided by Durante (2011: 119), who argues for 29 autonomous language groups divided over 9 genetic units. The estimate encountered in Golluscio & Vidal (2009: 3-4) leans towards that of Durante in terms of number of languages and genetic groupings. In this thesis, I stick to the conservative estimate provided in Campbell (275-8), who argues for the presence of members of the following language families in the Chaco:

i. Guaykuruan
ii. Matacoan
iii. Enlhet-Enenlhet
iv. Zamucoan
v. Lule-Vilela
vi. Tupí-Guaranían
vii. Possibly several (near)-extinct and un(der)documented languages

The third group in the list above, the Enlhet-Enenlhet language family, will be the focus of this MA thesis.

1.2 Sanapaná and the Enlhet-Enenlhet language family

One of the (at least) six indigenous linguistic groupings found in the Chaco is the Enlhet-Enenlhet language family. In this section, I will discuss issues pertaining to the internal genetic classification of this language family, the sociolinguistic situations of its member languages, and the current state of linguistic research focussing on these languages.

1.2.1 Genetic relations

The Enlhet-Enenlhet language family, formerly known as *Lengua-Mascoy*, is one of the smaller linguistic stocks present in the Gran Chaco area. The only scholarly source on the topic, Unruh & Kalisch (2003), argues that this family is comprised of six languages, all spoken exclusively in the Paraguayan part of the Chaco: Enlhet (traditionally known as Lengua Norte), Enxet (traditionally Lengua Sur), Angaité, Sanapaná, Enenlhet (traditionally Toba-Maskoy), and Guaná.

This has not always been the case, however. In the past, certain authors assumed Enlhet and Enxet to be two dialects of the same language (for example, Brinton 1898: 190), while others did not consider Angaité and Sanapaná to be distinct languages (such as earlier versions of the Ethnologue, see Grimes 2003 qtd. in Unruh & Kalisch 2003: 220). Furthermore, a seventh language has at times been assumed to belong to this language family. Sušnik (1962), for instance, posited the existence of *Emok-Toba*, a mixed language which would have originated in communities where Enenlhet-Toba and Toba-Qom

(Guaykuruan) groups lived together. Unruh & Kalisch (2003: 220) dismiss this claim, however, and their listing of Enlhet-Enenlhet languages is today the most influential one.

On the subgroupings within this language family, even fewer trustworthy data are available. Glottolog (Hammarström, Forkel & Haspelmath 2017) groups Enlhet, Enxet, Angaité and Sanapaná together as the Lengua-Sanapaná-Angaité subgroup, which then on a higher level group together with Guaná and Enenlhet. Such a family tree would look as follows (figure 1).

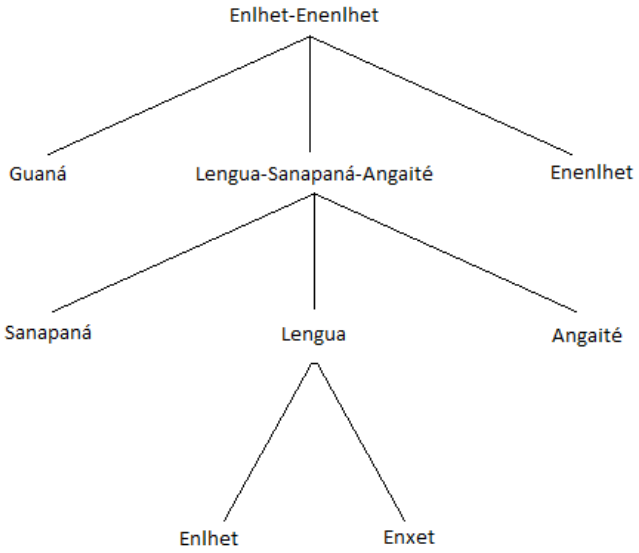

Figure 1: Genetic relationships amongst the Enlhet-Enenlhet languages according to Hammarström et al. (2017)

The evidence for this classification is rather scanty, however. The sole reference provided in its support is a brief comment in Gomes (2013: 72-3), who notes that Sanapaná, Angaité, Enlhet, and Enxet are more mutually understandable than either of them are with Enenlhet or Guaná. It must, however, be kept in mind that genetic relationships between languages are established on the basis of shared historical innovations rather than synchronic mutual intelligibility, since the latter can also be a result of borrowing and convergence through language contact in a later stage.

Unruh & Kalisch (2003: 210-3) do present some systematic data on the genetic classification of the Enlhet-Enenlhet languages, and come to a different conclusion, as represented in figure 2. They posit a Western and an Eastern branch of the family – the former consisting of Enlhet and Enxet, the latter of Sanapaná, Angaité, Enenlhet, and Guaná – based on a number of lexical (e.g. *sawow* vs. *peletaw* for 'knife') and grammatical innovations (e.g. *ak-* vs. *as-* as the first person singular cross-referencing prefix). Within the Eastern branch, a north-south division is posited between Guaná and Enenlhet on the one hand, and Angaité and Sanapaná on the other hand.

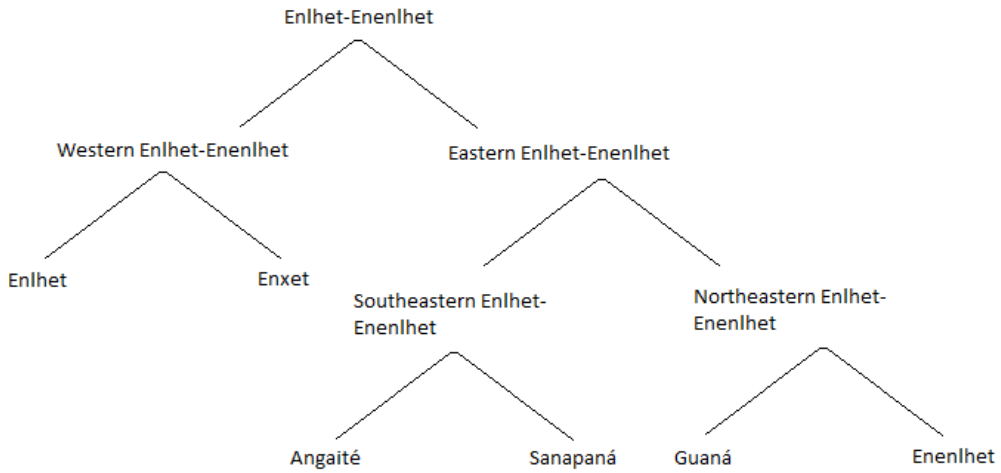

Figure 2: Genetic relationships amongst the Enlhet-Enenlhet languages according to Unruh & Kalisch (2003)

Nevertheless, Unruh & Kalisch (2003: 212) argue that the Enlhet-Enenlhet family should be seen as having evolved from a dialect continuum rather than from a number of neat binary splits in the proto-language. Within the Western branch, they argue that Enxet shares more features with the Eastern branch than Enlhet does, as such forming a sort of intermediate dialect. Similarly, within the Eastern branch, Guaná shows more similarities to Sanapaná and Angaité than Enenlhet does. The dialect continuum then can be represented as in figure 3, where circles group together the most closely related languages, but the lines between the different groups represent significant similarities between languages, including members of the different genetic subgroups (such as Enxet and Angaité, which Unruh & Kalisch show significant similarities despite belonging to different subbranches of the family).

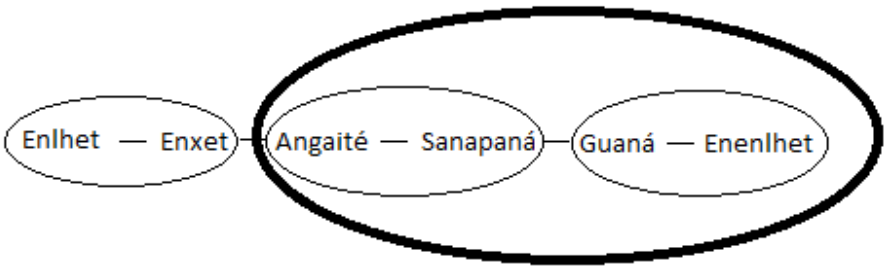

Figure 3: Proto-Enlhet-Enenlhet dialect continuum according to Unruh & Kalisch (2003)

This leads me to argue that any subsequent historical-comparative research into the Enlhet-Enenlhet languages should be conducted within the wave model of language change (advocated, amongst other, by François 2014). Since this methodology will be explained in more detail in subsection 2.3, suffice it to say here that this model is better suited than the traditional family tree model to explaining situations where language diversification occurred by means of a dialect continuum rather than clear and sudden divisions of populations. In particular, it allows for situations where some innovations affected, for example, only Enlhet and Enxet, whereas others affected Enxet, Angaité and Sanapaná. In this way, it would be able to provide an explanation for the gradient nature of the variation between Enlhet-Enenlhet languages.

1.2.2 Sociolinguistic contexts

In terms of the sociolinguistic situation, then, all Enlhet-Enenlhet languages can to some extent be called endangered. As mentioned before, the Chaco peoples have undergone a steady process of cultural assimilation to mainstream Paraguayan society over the last few centuries. For the Enlhet-Enenlhet groups, this started at the end of the 19th century with the annexation of the eastern Chaco by the Paraguayan state, and the establishment of the first tannin factories on the Paraguay river where indigenous groups were – both with and against their will – concentrated (Villagra Carrón 2010: 5). Here, many Enlhet-Enenlhet abandoned their heritage language in favour of the more economically useful lingua franca, Paraguayan Guaraní, and to a lesser extent Spanish. This process of language shift often continued in the missionary settlements where the Mennonite colonisers gathered the indigenous groups from the 1930s onwards (Villagra Carrón 2010: 6).

As a consequence, all Enlhet-Enenlhet languages have a rather low number of native speakers nowadays. For Guaná, the situation is the worst. According to Zanardini (2012: 3), only 14% of the ethnic group still speaks the ancestral language. Since this ethnic group consists of 393 people according to the Paraguayan census (DGEEC 2012), this comes down to around 50 Guaná speakers. According to Hannes Kalisch (p.c. 2017), the situation is even worse: he found only 4 living rememberers of the language, rendering Guaná effectively moribund. The most vital of these languages in terms of the numbers found in Zanardini (2012) and DGEEC (2012) is Enlhet Norte, with 96% of the ethnic population of 8100 still speaking the language, coming down to over 7500 native speakers. Sanapaná, the main language under discussion in this thesis, constitutes a middle ground, with around half of the almost 3000 ethnic Sanapaná still speaking the ancestral language.

Despite its comparatively low number of native speakers, however, Sanapaná is not a critically endangered language. There are two communities, La Esperanza and Anaconda, where virtually the whole population still speaks Sanapaná (88% and 93%, respectively, see DGEEC 2012). Additionally, it is in these two communities still transmitted to the youngest generations of children, who are often monolingual in Sanapaná until they enter school (personal observation), and it is used in almost all domains of daily life, including religious life (personal observation, Gomes 2013: 72).

1.2.3 Previous linguistic research

Descriptive linguistic research on the Enlhet-Enenlhet languages is very scarce. For two of the languages, Enxet Sur and Angaité, no descriptive work is available at all to my knowledge. Two others, Enenlhet and Guaná, are the subject of a didactic grammar (Unruh, Kalisch & Romero 2003 and Unruh & Kalisch 1999, respectively). Even though these books contain valuable data and interpretation, their linguistic use is rather limited, since they consist mostly of unglossed example sentences accompanied by a free translation in Spanish. Enlhet Norte, then, is documented in a grammar sketch in Susnik's (1977) ethnographical work, and Kalisch (2009) discusses the nature of arguments and predicates in the language.

For Sanapaná, ultimately, Gomes (2009a, 2009b, 2011, 2013) is the only linguist who published data and insights. These findings are, however, preliminary. In his earlier articles, he presents a sketch of the (morpho)phonology (Gomes 2009a,b) and the morphological mechanisms related to gender and number (Gomes 2011). In his doctoral dissertation (Gomes 2013), he again by and large focusses on (morpho)phonology and basic nominal and verbal morphology, with some forays into the syntax of simple

declarative, interrogative, and imperative clauses. As he indicates himself, this work is still in need of further investigation.

1.3 Research design and structure of the thesis

Before closing off this introductory chapter, I dedicate some attention in this section to a number of practical matters of interest to the reader of this thesis. In particular, I discuss the nature of the data used throughout the thesis and the way in which they were collected, I present the research objectives of this study, and present the reader with an overview of the structure of this work.

1.3.1 Data collection

This MA thesis aims to be a first addition to the corpus of descriptive linguistic work on Sanapaná. The data used throughout the thesis were gathered during a fieldwork trip to the Sanapaná community of La Esperanza, Departamento Presidente Hayes, Paraguay. I started planning the field trip in the fall of 2016. Via my supervisor Mily Crevels, Hannes Kalisch from *Nengvaanemkeskama Nempayvaam Enlhet*, and the Paraguayan NGO *Pro Comunidades Indígenas*, I managed to contact don Marino Ortega, the leader of the La Esperanza community. He generously agreed to welcome me in the community, where I eventually arrived early December 2016. After being offered an empty house in the community to stay in, about two kilometres from the community centre, I lived here for a week in order to get acquainted with the people of the community, discuss the ways in which we could work together, and allow enough time for people to ask questions about the nature and goal of my stay in the community.

Name	Age	Gender
Valerio Chávez	23	M
Regino Echeverría	33	M
Silio Recalde	65	M
Doto Teytaro	57	M
Cecilio Teytaro	47	M
Civito Montes	44	M
Marino Ortega	56	M
Esteban López	41	M
Cristino Benítez	33	M
Pisoa Cabrera	43	M
Roberto Álvarez	60+	M
Enrique Loma	51	M
Venancio González	60+	M

Table 1: Language teachers and storytellers during my 2016-2017 fieldwork in La Esperanza

After about a week, the community gave me their consent to start recording and working on the Sanapaná language. From this moment onwards, I did linguistic work in La Esperanza for a total of about six weeks. Altogether, I worked with 10 language teachers between the ages of 23 and 65, and a few other people who sporadically contributed short narratives for me to record (listed in table 1 above). These

were all male, because of the relatively low number of women in the La Esperanza community: it is more common for women then for men to move out of the community, after marriage or for employment on Mennonite farms, for example. Data were collected both through elicitation with these language teachers (using Spanish as a working language), and by transcribing and translating, with their help, the stories I recorded into Spanish. Throughout this thesis, examples from my own fieldwork will always be indicated in the form (FN: X.Y) immediately after the free translation, where the FN indicates that the example comes from my personal field notes, X stands for the notebook in which the example is written down, and the Y for the page number.

Example sentences from Enlhet-Enenlhet languages other than Sanapaná stem from the work of other authors. Examples from Enenlhet throughout this thesis are all taken from Unruh et al. (2003), Guaná examples are from Unruh & Kalisch (1999), and Enlhet examples are from Kalisch (2009). Examples from Angaité come both from Unruh & Kalisch (2003) and from Villagra Carrón (p.c. 2017), those from Enxet are from Unruh & Kalisch (2003) and Elliot (p.c. 2017). The spelling has been adapted to the conventions of this thesis (glottal stop represented by *ʔ* rather than ', lateral fricative by *ɬ* rather than *lh*, palatal glide by *j* rather than *i* or *y*, velar glide by *w* rather than *v*, velar nasal *ŋ* rather than *ng*, and vowel lengthening by the length mark *ː* rather than two subsequent vowel graphemes). As mentioned in the introduction, example sententences in the first two works are not glossed, but provided with a free translation in Spanish. The Spanish translations are translated into English by me, and I constructed the glosses to the best of my ability based on the grammatical notes provided in each chapter of Unruh et al. (2003). I use the terminology proposed by these authors throughouth this thesis, except when referring to Sanapaná. Examples from Kalisch (2009) are taken over with the original glossing, with minor changes for consistency throughout this thesis. Any errors resulting from this process are of course my own.

1.3.2 Research objectives

Because of the limited quantity of data at my disposal, it is not possible at this stage to provide a complete grammatical sketch of the language. I have instead chosen to focus on one specific grammatical feature: the discourse particles *łke*, *ałta*, and *hata* I will describe both their phonological form and morphophonetic processes in which they take part; their syntax, exemplifying the way in which they combine with other parts of speech; and their semantics, clarifying which functions they serve in Sanapaná clauses.

In particular, I argue that all three of them have TAM functions: *łke* refers to an event or state of affairs which took place earlier on the day of the utterance, or is just about to take place (making it an immediate aspect marker); *ałta* refers to events taking place in the past before the day of the utterance (making it a prehodiernal past marker); and *hata* refers to events in the non-immediate future. Often, these particles combine syntactically with verbs: they then act to specify the time reference of the event expressed by the verb, since the Sanapaná verbal TAM system, in my analysis, does not encode tense overtly on the verb or in auxiliaries. Nevertheless, they co-occur with other word classes as well, such as nouns and conjunctions. In these cases, they trigger a predicative reading of the word they follow, overtly expressing the predicativity that is inherent in these other word classes (making Sanapaná, as is Enlhet Norte according to Kalisch 2009, an omnipredicative language). As such, I argue that these particles fulfil both the function of temporal reference and that of predication. This thesis then aims to contribute both to the descriptive literature on Sanapaná, and hopes to provide data for the typological study of these two features.

The third research objective of this thesis consists in presenting a comparison of the Sanapaná data I collected with the available data on Enlhet (Kalisch 2009), Guaná (Unruh & Kalisch 1999), and Enenlhet (Unruh et al. 2003). By examining, inasmuch as this is possible, the formal and functional characteristics of these three particles in other languages of the family, I hope to contribute to the investigation of the genetic relations between the Enlhet-Enenlhet languages. My findings generally support Unruh & Kalisch' (2003) idea that the Enlhet-Enenlhet languages developed from a dialect continuum: no neat splits between (groups of) languages are found. Instead, formal and functional features of these particles often overlap in (groups of) languages. It will be argued here that Guaná, Enenlhet and Sanapaná share most features to the exclusion of Enlhet, a finding in agreement with Unruh & Kalisch (2003), even though Enlhet shows a number of similarities to Guaná and (especially) Enenlhet as well. Within the Eastern Enlhet-Enenlhet branch, then, my findings contradict those of Unruh & Kalisch (2003): Sanapaná and Guaná seem to share more characteristics than Guaná and Enenlhet.

1.3.3 Structure of this thesis

The next chapter of this thesis provides the theoretical background necessary for my subsequent analysis. It contains a literature review on the topics of tense-aspect-mood phenomena, both related to verbs and to nouns, omnipredicativity, and the wave model of language diversification applied in the comparative part of this study. In chapter 3, then, I give an introduction to Sanapaná grammar. Here, I discuss the typological profile of the language, the phonology and orthography, the verbal cross-referencing system in transitive and intransitive clauses, verbal manifestations of tense, aspect and mood, and other aspects of Sanapaná grammar that are necessary to the understanding of the discussion in the next chapter. I introduce the particles under scrutiny in chapter 4, analysing them both formally (in terms of their morphology and syntax) and functionally (in terms of their semantics). In chapter 5 I link these empirical data back to the theoretical foundations laid in chapter 2, discussing them in terms of the proposed frameworks on TAM and flexible word classes. Chapter 6 broadens the scope of my investigation and compares the behaviour of Sanapaná *łke*, *ałta*, and *hata* with that of their cognates in Guaná, Enlhet Norte and Enenlhet, the only three other Enlhet-Enenlhet languages for which enough data is available. Chapter 7, ultimately, summarises and concludes this thesis, presenting a number of suggestions for further research.

Chapter 2 Theoretical Background

Before discussing the Sanapaná data concerning the particles *łkek*, *ałta*, and *hata*, I first present some theoretical background necessary to analyse their functions of temporal reference and predication. In particular, the first section of this chapter will provide a literature overview of research on tense, aspect and mood (TAM) categories. The second section, then, will summarise the relevant literature on word class flexibility. Throughout these two sections, I will exemplify theoretical concepts both with examples given in the literature discussed, and with relevant examples from the Enlhet-Enenlhet languages. The third and last section will briefly present the wave model of language diversification which will later in chapter 6 be used for the comparison of the Enlhet-Enenlhet languages.

2.1 Tense-Aspect-Mood categories

Tense, aspect and mood are three categories that have been the subject of much discussion over the years, both from an empirical-descriptive point of view (describing the TAM systems of individual languages), and from a theoretical-typological perspective. They are usually seen as closely intertwined verbal categories in most languages (see, for instance, Rathert & Musan 2012), although a lively debate on nominal tense was held relatively recently as well (amongst others in Alexiadou 2008; Nordlinger & Sadler 2004a,b; Tonhauser 2007, 2008).

2.1.1 Verbal tense

Verbal tense can be defined as "an inflectional marker on the verb used for denoting the *temporal location* of an event (or situation)" (Bhat 1999: 13, emphasis original). This temporal location must always be specified with respect to another event: either the moment of speaking, or some other event specified in the utterance. For several decades, and arguably until today, Reichenbach's (1947) model was the standard theoretical model in which to describe tense phenomena in different languages. In this model, taken over by amongst others Bhat (1999) and Stowell (2012), tense is the expression of the relationship between three moments in time: speech time (S, the moment when the utterance is produced), event time (E, the time or time interval at which the event expressed in the utterance took place) and reference time (R, the time in relation to which the event expressed is situated).

Any two of these three moments in time can either be simultaneous, or one can precede the other. To give an example from English: in a simple past clause such as (1a), S and R coincide and E precedes

both of them (i.e. E, the *walking* event, is located with respect to the time of the utterance). The timeline in such a clause can then be represented as follows: E > R/S In the English past perfect *had left* (1b), on the other hand, R precedes S, and E is situated before R: the *leaving* event takes place before the *arriving* event, which in turn precedes the speech time. Here the timeline looks different: E > R > S.

(1) English (Indo-European, inspired by Bhat 1999)
   a. I walked to school.
   b. When I arrived, he had already left.
   c. I will walk to school.
   d. I will have left when he arrives.

The same can be done when the event expressed in the utterance follows the speech act. In (1c), for example, R and S coincide again, and both are followed by E: the *walking* will take place after the sentence is uttered (R/S > E). Similarly, for the future perfect *will have left* in (1d), R follows S (the *arriving* will take place after the speech time), and E precedes R but follows S (the *leaving* event is yet to take place at the speech time, but will occur before the *arriving* event). Here, the timeline is S > E > R. In the Enlhet-Enenlhet languages, clauses expressing a temporal situation where R and S precede E typically take the form of example (2). Which components of the temporal meaning of these sentences are carried by which constituents will be treated in chapter 3.

(2) Enenlhet (Enlhet-Enenlhet, Unruh et al. 2003: 104)
   popiet *ko* *s-makʔ* *a-teŋj-akʔ*
   deer PRO:1SG 1SG-want 1SG-search-FUT
   'I am going to search for deer.'

Tenses such as those found in (1a) and (1c), where the event time is in effect located with respect to speech time, can be called *deictic tenses*: they have a deictic reference point in the speech act. The tenses found in (1b) and (1d), alternatively, can be called *non-deictic tenses*. Deictic tenses can maximally show a threefold distinction: they can refer to an event prior to the speech act, simultaneous with it, or posterior to it. Non-deictic tenses, on the other hand, can potentially show nine different values: prior to, simultaneous with, or posterior to a reference time which is in turn prior to, simultaneous/overlapping with, or posterior to the speech time (Bhat 1999: 15-6).

An additional parameter that languages can use in their tense systems to make the temporal reference of an utterance more specific is that of remoteness. The most common remoteness distinction, in De Haan's view (2011: 450-2), is that between a hodiernal and a prehodiernal past, i.e. a deictic past tense to refer to events which happened earlier on the day of the utterance, and one to refer to events which happened at any time before that. Nevertheless, more complex systems are certainly possible, with the Yimas language of Papua New Guinea, for instance, distinguishing between two future tenses (hodiernal and post-hodiernal), a present tense (here, remoteness distinctions are for obvious reasons impossible), and three past tenses (a yesterday-past, one for the period between the day before yesterday and about five days ago, and a remote one, Foley 1991: 236).

(3) English (Indo-European, inspired by Bhat 1999)
   a. I will write a book.
   b. I will write a book tomorrow.
   c. *I wrote a book tomorrow.

(4) Sanapaná (FN2.113)
   a. *o-tjen-e=ta*
      1SG-sleep-IRR=FUT
      'I am going to sleep.'

b. *o-tjen-e=ta* *sosokha*
1SG-sleep-IRR=FUT tomorrow
'I am going to sleep tomorrow.'

c. **o-tjen-e=ta* *aknem=ałta*
1SG-sleep-IRR=FUT sun/day=PHOD
'I am going to sleep yesterday.'

Furthermore, it must be kept in mind that tense morphology on the verb is not the only way in which languages achieve temporal reference. Temporal adverbials, in Bhat's (1999: 35) terms, "[establish] a parallel structure that is related to the temporal structure that the tense system denotes." These temporal adverbials depend upon verbal tense (in languages that have it), and at the same time expand upon it: a clause such as (3b) with an adverbial *tomorrow* has more specific temporal reference than (3a), where the general future tense is used. Nevertheless, this adverbial depends on verbal tense mechanisms, since it cannot be used in clauses where it is incompatible with the temporal reference expressed by the verbal tense (3c). The same principle holds in Sanapaná: (4b) is grammatical and makes the temporal meaning of (4a) more specific since the semantics of the adverbial 'tomorrow' are compatible with the future semantics of the verb. On the other hand, (4c) is ungrammatical since the meaning of the adverbial clashes with that expressed on the verb.

2.1.2 Nominal tense

Although the default position in the past was that tense is a grammatical category inherently linked to the verb, and hardly compatible with the noun (Bybee 1985: 161; Pinker & Bloom 1990: 715), over the last two decades a lively debate about nominal tense has taken wing. In particular, Nordlinger & Sadler (2004a, b) have been proponents of the existence of nominal tense. These authors identify two distinct kinds of temporal reference that can be expressed morphologically on the noun: independent nominal tense and propositional nominal tense.

The former kind of nominal tense is found, in their view, most often and is especially common in languages of the Americas (Nordlinger & Sadler 2004a: 779). This type of nominal tense has scope only over the noun it occurs on: it "locates the time at which the property denoted by the nominal holds of the referent or, in the case of possessive phrases, the time at which the possessive relation holds" (Nordlinger & Sadler 2004a: 779) As such, it is completely independent from the clausal TAM. Independent nominal tense can be exemplified by means of the Guaraní sentences in (5a-b). Arguably, the Sanapaná adverbial *aknem=ałta* (as seen before in 4c) is another example of this: the past marker *ałta* locates the referent of the noun *aknem* 'day' in the past.

(5) Paraguayan Guaraní (Tupí-Guaranían, Gregores & Suárez 1967, via Nordlinger & Sadler 2004a: 781)

a. *o-va-ta* *che-róga-kue-pe*
3-move-FUT 1SG-house-PST-in
'He will move into my former house.'

b. *a-va-va'ekue* *hóga-rã-pe*
1SG-move-PST 3.house-FUT-in
'I moved into his future house.'

In (5a), the past morpheme *-kue* locates the nominal *róga* 'house' in the past. More specifically, since the house is possessed, it locates the time at which the possessive relationship between the house and the first person owner was valid in the past: 'the house that used to be mine, but is not anymore.' The

temporal location of the nominal is independent of that of the clause as a whole, however: the time of the event represented in this clause, the 'moving in,' is located in the future by the *-ta* suffix on the verb. The reverse is true for (5b): the possessed nominal is located in the future ('the house that will belong to him') by the *-rã* suffix, whereas the verbal tense suffix locates the proposition as a whole in the past.

On the other hand, some languages show what Nordlinger & Sadler (2004a, b) call "propositional nominal tense," where the temporal reference morphologically expressed on the noun phrase does have scope over the whole clause. In Chamicuro, for example, the past-nonpast tense distinction is usually not expressed on the verb, but on the definite article. Even though it is phonetically encliticised to the verb, Parker (1999: 553) and Nordlinger & Sadler (2004b: 598) argue that this article is not morphologically part of the verb, but still a constituent of the NP. Examples (6a-b) show that the use of *na* or *ka* as definite article locates the state of affairs expressed in the clause in the present or in the past, respectively.

(6) Chamicuro (Arawakan, Parker 1999: 533, via Nordlinger & Sadler 2004b: 598)

a. *p-aškalaʔt-ís=na* *čamálo*
2-kill-2PL=DET(NPST) bat
'You (PL) are killing the bat.'

b. *p-aškalaʔt-ís=ka* *čamálo*
2-kill-2PL=DET(PST) bat
'You (PL) killed the bat.'

A last use of nominal tense, in Nordlinger & Sadler's (2004a,b) view, is that it can also interact with discourse. In Somali, for example, tense marked on nominals can not only locate the existence or possession of a referent in the present or past, but also its introduction into the conversation (Nordlinger & Sadler 2004a: 787). In (7a-b) it can be seen how Somali definite determiners (*-da* and *-dii* suffixed to the nominal) can express independent nominal tense (locating the *crisis* in the present or the past, respectively).

(7) Somali (Cushitic, Lecarme 1999: 335, via Nordlinger & Sadler 2004a: 786-7)

a. *dhibaatá-da* *Khalíij-ku* *welí* *way* *taagán* *tahay*
problem-DET.F gulf-DET.M.NOM still FOC.3S permanent is
'The crisis of the Gulf still persists.'

b. *dhibaatá-dii* *Khalíij-ku* *wáy* *dhammaatay*
problem-DET.F.PST gulf-DET.M.NOM FOC.3S end.PST
'The crisis of the Gulf ended.'

c. *ardáy-da* *baan* *kasin* *su'áash-aadii*
students-DET.F FOC.NEG understand.PST question-DET.F.POSS2S.PST
'The students (who are present/I am introducing now) did not understand your question.'

d. *ardáy-dii* *way* *jogaan*
students-DET.F.PST FOC.3O be.present.NPST
'The students (whom I introduced to you before) are present.'

In (7c-d), on the other hand, the use of the definite determiners *-da* and *-dii* with the noun *ardáy* 'students' does not give any information concerning whether the referents of this nominal were students in the past or are students in the present. Instead, it indicates when the referents that are designated as 'students' were introduced in the discourse. A similar phenomenon can be perceived in the Enlhet sentences in (8a-b): the form *ałta* is used in (8a) to locate the time at which the referent of *majkaʔa* possessed the quality of being a 'visitor' in the past. In (8b), however, it locates the time at which the referent of *semheŋ* 'dog' was introduced into the discourse context in the past.

(8) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 123; 134)

a. *e-l-pejwe:s-a-mk-eʔ* *ałta* *enłet* *majkaʔa* *ałta* *kwesejʔ*
1SG-DISTR-greet-FACT-CPL-PRIM[1] PHOD man/enlhet visitor PHOD then
'The people greeted me, I was a visitor then.'

b. *k-etsep-kek* *nek* *semheŋ* *ałta* *k-jew-ejʔ*
2/3F-die-FACT.PRIM RPT dog(F) PHOD F-become_big-INF
'They say that the large dog we saw before died.'

From a theoretical point of view, Nordlinger & Sadler (2004a: 778) put four requirements which a morpheme must fulfil in order to be considered a nominal tense marker. The morpheme must:

i. express a distinction in a category of tense, where tense is defined as it would be for verbs
ii. be productive across the whole word class
iii. not be restricted to nominals functioning as predicates of verbless clauses, but also be encoded on NPs in clauses headed by verbs
iv. be a morphological category of the nominal word class.

Tonhauser (2007; 2008) has a different way of describing nominal tense: in her view, nominal temporal marking expresses the relation between a noun phrase time (t_{NP}), which corresponds to the R of verbal tense, and a nominal time (t_{NOM}) which corresponds to the E of verbal tense. Furthermore, she specifies the first criterion of the four listed above (Tonhauser 2007: 862; 2008: 333). In order for a set of nominal markers to meet the definition of tense in the same way as this would be defined for verbs, they must show the following four extra characteristics:

i. Their meaning must not inherently encode a state change
ii. It must be possible for their meaning to be "contextually determined by anaphorically resolving the reference time to a contextually given time" (Tonhauser 2007: 860), as in English *"Sheila had a party last Friday and Sam got drunk*" (Tonhauser 2007: 860), where the time of Sam's getting drunk is anaphorically related to the time of Sheila's giving a party.
iii. Their use must be mutually exclusive, i.e. two of them cannot co-occur.
iv. It must be possible for a temporal modifier (e.g. an adverbial) to constrain their reference.

These eight criteria are the ones which will in the rest of this thesis be used to determine whether or not Sanapaná and the Enlhet-Enenlhet languages can be said to show nominal tense. This point being made, I move on to the discussion of aspect.

2.1.3 Aspect

While tense, as explained before, is usually seen as a morphological mechanism to express the temporal location of an event, aspect is typically defined as a (verbal) category which serves to express the temporal structure of an event (Bhat 1999: 43), or in other words, "*how* an event unfolds" (De Haan 2011: 453, emphasis mine) rather than when it does.

[1] The complexive is, according to Unruh et al. (2003: 166), an aspectual category of the Enlhet-Enenlhet languages which is used to represent an event as consisting of various subevents which form a unit together. The primative (what I call V1-form) of a verb is the form used when this verb occurs in its non-marked position, i.e. clause-initially. Unruh et al.'s (2003) factive roughly corresponds to my realis in Sanapaná.

According to certain authors (for instance, Sasse 1991, 2006), two types of aspect must be distinguished. On the one hand, there is Aktionsart, which has also been designated with the terms "lexical aspect" or "situational aspect" (Bhat 1999: 45). This term is used to refer to different types of situations that are expressed by verbal lexical items, regardless of any aspectual morphology (such as events, processes, states, etc.). In English, for example, it can be said that *hit* and *walk* are verbs with a different Aktionsart. The lexical meaning of the former implies a single momentary event, whereas that of the latter inherently involves duration.

On the other hand, there is viewpoint aspect or grammatical aspect. According to authors such as and Bybee (1985) and Siewierska (1991), the semantics of viewpoint aspect can be divided up into three distinct domains. Firstly, there is the domain of perfective as opposed to imperfective aspect: the opposition between them is, according to Bybee (1985: 141), the most common aspectual distinction found cross-linguistically. Traditionally, these categories are said to reflect whether an external viewpoint towards the event in question is taken on, portraying it as a unit (perfective); or an internal viewpoint, portraying it as ongoing and incomplete (imperfective).

In the definition of perfective and imperfective aspect, one can make use of Reichenbach's (1947) E and R as well (S is less relevant here, see Siewierska 1991: 117), which demonstrates the intricate relationship between the categories of tense and aspect. If an event (taking place at an event time E) is to be portrayed as a completed whole, the whole of E must necessarily be included within the reference time R. This is exemplified in (9a): here, the event time is the period of time it took the producer of this sentence to read Milosz' poetry. The reference time with respect to which this E is situated is the whole period of this speaker's youth. In order to use the perfective aspect to present the poetry reading as one whole, completed event, as is done in this sentence, it is necessary that all of the poetry-reading took place in the speaker's youth.

(9) Polish (Indo-European, Siewierska 1991: 117, adapted from Majewicz 1985: 79)

a. *W* *młodości* *przeczytałam* *wiersze* *Miłosza*
in youth PFV.read:PST.3SG.F poems Miłosz
'In my youth I read (all of) Milosz' poetry.'

b. *W* *młodości* *czytałam* *wiersze* *Miłosza*
in youth read:IPFV:PST.3SG.F poems Miłosz
'In my youth I read Milosz' poetry.'

In order for the imperfective aspect to appear, on the other hand, the opposite situation must hold true. The event time must extend beyond the reference time with respect to which it is portrayed in order to be able to represent the state of affairs as ongoing or incompleted. This is the case in (9b). Here, it is implied that the speaker's youth is over (i.e. R is completely in the past), whereas he is still reading Milosz' poetry (i.e. E started in the past and goes on until now, and is consequently not completely contained within R). Therefore, the imperfective aspect can be used felicitously in this clause.

The second domain within viewpoint aspect is that of phasal aspect, which is used to "emphasize different phases of the development of [a state of affairs] through time" (Siewierska 1991: 118). According to Siewierska (1991: 118), languages can distinguish, for example, prospective, ingressive, progressive, egressive, and/or resultative aspect. Verb forms with morphology expressing these aspectual categories emphasise, respectively, the moment immediately before the start of an event, the event's beginning, the event in progress, the final phase of an event, and the moment immediately after the event.

A language which makes use of at least two of these distinctions is Mocoví, another language spoken in the Gran Chaco, which belongs to the Guaykuruan family (Carrió 2010). It overtly marks progressive

and prospective aspect, exemplified in (10a-b). In example (10a), E (the period in which the subject is searching wood) started before R (which in this case coincides with S), and goes on after reference time. As such, the event is portrayed as ongoing at R, and the progressive aspect is used. In (10b), then, the prospective aspect is used: the R with respect to which the *speaking* event is located (the *now*, which again coincides with speech time) occurs immediately before the start of the E itself (i.e. the event is *about to happen* at R).

(10) Mocoví (Guaykuruan, Carrió 2010: 247-8)

a. *s-anip-tak*
1SG-search.wood-PROG
'I am searching wood.'

b. *n-aʔGa-aG-ontiño* *nagi* *r-taqa-o* *xuan*
IND-listen-1PL-? ADV:TEMP 3SG-speak-PROSP Juan
'We are listening now to Juan who is about to speak.'

The third and last domain of viewpoint aspect is *quantificational aspect* (see e.g. Siewierska 1991: 118 and Bhat 1999: 44). As the label gives away, quantificational aspect gives information concerning the number of (sub)events involved, or the frequency with which an event occurs. Here too, several distinctions can be marked. Semelfactive aspect, for instance, is used to refer to unique events which only occur once.

Habitual aspect, then, expresses events whose repetitive nature is established inductively. An example given by Bhat (1999: 53) is that of the arrival of a train: one does not need to deduct the habit of arriving at a certain time from multiple observations of this arrival, since one can generate it inductively from a rule (in this case, a timetable).

(11) Hindi (Indo-European, Bhat 1999: 55)
vah niyamit ru:p se yahã: a:y-a: *kar-ta:* *tha:*
he regular way by here come-PFV do-IPFV was
'He used to come here regularly.'

For the frequentative aspect, on the other hand, multiple observations of the event are necessary to make a generalisation. Therefore (11), with the frequentative construction consisting of the perfective form of the main verb followed by the imperfective form of *do*, must always be based on multiple observations of the *coming* event.

Iterative aspect, ultimately, like frequentative aspect, refers according to Bhat (1999) to the occurrence of multiple (empirically witnessed) events. The distinction between the two lies in the fact that the former is used to express repetition of an action or event on the same occasion (i.e. it presents an event as consisting of a number of identical sub-events), whereas the latter "portrays events repeated on different occasions" (Bhat 1999: 53). For example, iterative aspect could be used in a sentence conveying the meaning *he scratched his back*, where the speaker witnessed multiple scratching motions on one specific occasion, whereas frequentative aspect is more suited to sentences like (11) where a significant amount of time elapses between the occurrences of the event. The Enenlhet aspectual category of *repetitive* (Unruh et al. 2003: 78) can also be seen as a category of quantitative aspect: it is used to mark single occurences of events that take place frequently and habitually. In (12), for example, the repetitive marker *-kha-* indicates that the sleeping event to which this sentence refers, is one instance of a habitual pattern of sleeping.

(12) Enenlhet (Enlhet-Enenlhet, Unruh et al. 2003: 74)
maʔ a-tjene-kha-k koʔo
want 1SG-sleep-REP-FUT PRO:1SG
'I am going to sleep again.'

2.1.4 Mood

The last category to be discussed in this section is that of mood. Whereas tense and aspect respectively express the temporal location and temporal structure of an event, mood rather expresses different components of its actuality, or its truth value (Bhat 1999: 63). Just like aspect and tense, mood can be described in terms of various parameters (see, amongst others, Bybee 1985; Bhat 1999; De Haan 2005): the speaker's judgment towards an event, the kind of evidence available for a proposition (both subsumed under the term epistemic or knowledge-based mood), and the conditions which influence an event to take place (deontic or action-based mood).

With respect to the first parameter, a speaker can judge an event to be amongst others, real, unreal, certain, probable, or improbable. A common distinction made by languages with regards to this judgment, is the realis-irrealis distinction (De Haan 2005: 43). The former refers to events that are "actualised or […] actually occurring" (Bhat 1999: 65), whereas the latter did not occur in the real world (yet).

One language showing such a distinction is Chalcatongo Mixtec: here, verbs have two distinct stems, one termed *realis*, used for events which are habitual, in progress, or completed, and one *potential*, used for events that are possible or probable, but non-actualised (including for example, events that are expected to take place in the future, conditionals and counterfactuals, Macaulay 1996: 45-6).

Sanapaná (and the other Enlhet-Enenlhet languages) also arguably demonstrates such a realis-irrealis distinction, as illustrated by the two forms of the verb for 'sleep' in (13a-b). The first form, with the suffix *-ek*, is used to express events which are portrayed as hypothetical or as yet unrealised, whereas the form with the *-eje* suffix is used to represent events that have been or are in the process of being actualised. This distinction is also accompanied by the use of different sets of pronominal argument prefixes. Across the Enlhet-Enenlhet language family, these forms in *-ek* typically convey meanings related to desire and future temporal reference (Unruh et al. 2003: 46), or, in other words, non-realised events.

(13) Sanapaná (FN2.113)
a. *o-jetn-ek*
1SG-sleep-IRR
'I am going to/will sleep.'

b. *as-jetn-eje*
1SG-sleep-REAL.V1
'I am sleeping/I slept.'

As can already be gleamed from the translations of sentences (13a-b), mood is closely interlinked with tense and aspect as well. On a certain level, the modal irrealis-realis distinction can be seen to semantically map onto the temporal future-nonfuture distinction: events can, arguably, only be judged as *actual* if they have already occurred or are in the process of occurring (as in 13a), whereas events that are to happen in the future are by definition not actualised (such as 13b). Why it is, then, that this distinction should be seen as a modal realis-irrealis distinction rather than a temporal future-nonfuture one will be clarified in section 3.4.

The second category of epistemic modality, then, has to do with the evidence a speaker has for a certain utterance. Through evidential marking, a speaker can, amongst others, indicate whether he acquired the information expressed in his utterance by seeing or hearing the event happening, being told about it, or inferring it from indirect evidence (Bhat 1999: 64). This is done, for example, in Tuyuca (Barnes 1984: 257). The example sentences in (14a-e) show respectively the suffixes marking visual evidence, aural evidence, inferred evidence, hearsay evidence, and assumed evidence. The Enlhet example in (8b), repeated here as (15), can also be said to show such an evidential marker: the *nek* indicates that the speaker has hearsay (i.e. reported) evidence for the proposition.

(14) Tuyuca (Tucanoan, Barnes 1984: 257, via Bhat 1999: 71)

a. *díiga apé-wi*
'He played soccer (I saw him play)'

b. *díiga apé-ti*
'He played soccer (I heard soccer being played, and I heard him playing)'

c. *díiga apé-yi*
'He played soccer (I saw indirect evidence of him playing, e.g. his dirty shirt)'

d. *díiga apé-yigɨ*
'He played soccer (someone told me so)'

e. *díiga apé-hĩ̃yi*
'He played soccer (it is reasonable to assume so, e.g. he always plays this time of day)'

(15) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 134)

k-etsep-kek	*nek*	*semheŋ*	*ałta*	*k-jew-ej?*
2/3-die-FACT.PRIM	RPT	dog(F)	PHOD	F-become_big-INF

'They say that the large dog we saw before died.'

A last note with respect to evidentiality is that not all scholars agree that it should be subsumed under the category of modality. The reason why evidentiality is treated as epistemic modality in such works as Bybee (1985), Palmer (1986) and Frawley (1992), is that, according to these authors, the source one has for a certain piece of information determines in part one's veridicality judgment towards it: a statement for which one has seen the evidence with one's own eyes is more likely to be judged true or certain (and as such to be seen as realis) than a statement for which one only has indirect evidence (De Haan 2005: 49).

De Haan (2005: 50-51), on the other hand, argues that such an analysis is only achieved by forcing a (false) correspondence with English epistemic modals upon, for example, Tuyuca sentences. Instead, he proposes to analyse evidentiality as a deictic category, denoting the distance between a speaker and the proposition. In my view, these arguments are not necessarily contradictory: evidential categories can be both deictic and modal, just as, for example, a simple past can be both deictic and a tense. The aforementioned solution, to treat judgments and evidentials as distinct components of epistemic modality, is the one that will be adopted where relevant in the present thesis.

The last subdomain of modality that will be treated here is that of deontic modality, which expresses "the kind of compulsion which makes it necessary for an event to take place" (Bhat 1999: 75). As such, it is, according to Bhat (1999: 75), comprised of notions such as ability, willingness and desire (internal compulsions), and necessity, request and order (external compulsions). Just like evidentiality, deontic modality is linked to veridicality judgments: the stronger the compulsion for an event to take place (desire, for instance, is a stronger compulsion than mere willingness), the more certain a speaker can be that the event was/will be actualised.

One can easily exemplify deontic modality with English phrases, since English has a relatively extensive system of deontic modal auxiliaries. Take, for example, the sentences in (16a-c), which only differ in the choice of modal auxiliary. *May* in (16a) expresses permission, i.e. a weak, external compulsion for the *you* to perform the action expressed through the main verb. *Should* in (16b), on the other hand, expresses necessity, a stronger form of external compulsion, and *have to* in (16c) expresses an obligation, one of the strongest forms of external compulsion.

(16) English (Indo-European, inspired by Bhat 1999)
   a. You may leave now.
   b. You should leave now.
   c. You have to leave now.

The close link between epistemic and deontic modality can also be explained by means of English examples. In (16a-b), for example, the modal auxiliaries are ambiguous between an epistemic or a deontic reading. Without any further context, *may* in (17a) can express either permission (i.e. weak deontic modality: "he is allowed to come tomorrow") or possibility (i.e. weak epistemic modality: "it is possible that he comes tomorrow"). Similarly, *should* in (17b) can express the stronger epistemic notion of probability (i.e. "the book is most likely on the shelf"), or the stronger deontic notion of necessity (i.e. "the proper place for the book is on the shelf").

(17) English (Indo-European, Bhat 1999: 76)
   a. He may come tomorrow.
   b. The book should be on the shelf.

This is, however, not the case in every language. Ladakhi (Tibeto-Burman), for example, has a set of deontic modal suffixes (expressing, amongst others, ability, desire, and permission), which are clearly distinct from the epistemic modal suffixes (Koshal 1979: 228 via Bhat 1999: 76). According to van der Auwera and Ammann (2005), non-overlap between the structures used to express epistemic and deontic modality is found in most languages of the world. Therefore, it is warranted to keep them separate as two distinct but related subfields of modality.

Before moving on to the second section of this chapter, I wish to draw attention to the fact that temporal, aspectual, and modal categories might not always be as neatly distinguishable as it might seem from the examples and discussion above. According to Dahl (1985: 9), it is perfectly possible for a TAM category in a language to contain both temporal and aspectual meaning, for example. According to Deo (2012: 159), it can even be said that future tenses inherently contain modal meaning. Even though, on the one hand, these categories can be interesting precisely because they do not neatly fit into either class, for comparative purposes this is often inconvenient (Dahl 1985: 20).

Therefore, one can usually investigate which parameter is most strongly present in the semantics of a given category, and take this as the basic meaning of the category. The category of perfective, for example, in many languages combines semantics of past time reference and perfective aspect. Usually, however, the perfective semantics are more salient than the past ones: it typically occurs more often with perfective semantics and non-past reference than with past time reference but imperfective semantics. Consequently, the parameter of perfectivity can be seen as dominant, and the category as a whole characterised as an aspectual category (Dahl 1985: 23). Such an approach will be followed throughout this thesis.

2.2 Word class flexibility

The notion of *predicate* has been in use for decades, as a central notion both in linguistics and in formal logic. In this section, I will first (subsection 2.2.1) clarify the meanings and uses of the term *predicate* and the related terms *argument* and *referent*, and of the functions associated with them (*predication* and *reference*). In subsection 2.2.2, then, I will give a brief overview on the literature on word class flexibility, discussing the ways in which languages can neutralise the distinction between word classes functioning as predicates and classes functioning as referents (the *omnipredicative* or *monocategorial* languages).

2.2.1 Predicates and referents

According to a standard dictionary (Merriam-Webster 2017), the term *predicate* is defined, in formal logic, as "something that is affirmed or denied of the subject in a proposition" or "a term designating a property or relation." The linguistic definition, according to the same dictionary, is "the part of a sentence or clause that expresses what is said of the subject and that usually consists of a verb with or without objects, complements, or adverbial modifiers." Both these definitions seem to have in common that they define as a predicate that part of a proposition which provides the information concerning what the subject does, what happens to the subject, or in what state the subject is. In formal terms, a predicate seems to be defined here as "the whole clause minus the subject." Following such a definition, in the English sentence (18a), *John* would be seen as the subject, and *gave Mary a book* would be seen as the predicate, i.e. what is being said about *John*.

Most recent linguistic investigations on the notion of predicativity (Evans & Osada 2005; Haude 2009; Hengeveld, Rijkhoff & Siewierska 2004; Launey 1994; Rijkhoff & Van Lier 2013; Queixalós 2006) have, however, made use of a narrower definition of the term *predicate*, which will be taken over in this thesis. According to this definition, stemming from Blake (1990) and Perlmutter (1980), the sentence is not split in a binary way between subject and predicate. Instead, a clause is taken to consist of a predicate and its (one or more) arguments. In a sentence such as (18b) below, this does boil down to the aforementioned subject-predicate split between the subject *John* and the predicate *run*, but in (18a), the predicate *give* has three arguments: its subject *John*, its direct object *the book* and its indirect object *Mary*.

(18) English (Indo-European)
   a. John gave a book to Mary
   b. John is running
   c. John was running yesterday.

In the rest of this thesis, propositions will thus be analysed as predicates which select a number of arguments. Additionally, constituents can appear which are not selected by the predicate itself: this is for example the case of the adverb *yesterday* in (18c). These constituents, termed *adjuncts*, do not have the same status as arguments: *yesterday* in (18c) can be omitted without making the clause ungrammatical, whereas *John* cannot.

With the distinction between predicates, arguments, and adjuncts defined for the purpose of this thesis, I turn to the functions associated with arguments and predicates. The former typically perform the function of *reference*, i.e. they refer to entities in the extralinguistic world which are involved in some way in the state of affairs expressed by the proposition. This is the case in the sentences in (18): the

arguments *John* and *Mary* establish reference to specific persons in the extralinguistic world who are known by these names and are involved in the *giving* and *running* events. Similarly, the argument *a book* in (18a) establishes reference to a specific book that was given to Mary by John, even though from the utterance it is not clear which exact book it is. Crosslinguistically, the reference function is typically associated with the grammatical category of nouns (and nominal phrases), since their semantics (denoting entities) are most compatible with this function (Queixalós 2006: 250).

The function of the predicate, then, is quite understandably called *predication*: it expresses the state of affairs in which the referents denoted by its arguments are involved. This can be an event (*give* in 18c) or a process (*run* in 18b), but also a state or a property. The function of predication is typically performed by verbs, as in the example sentences above, because of their semantic compatibility with the denotation of non-time-stable concepts such as events (Queixalós 2006: 250). Nevertheless, this is not always the case. In English, for example, nouns and adjectives can also function as predicates when accompanied by the copula *be*. In (19a), the nominal phrase *a teacher* functions as a predicate denoting a state in which its argument *John* is. Similarly, in (19b), the adjective *red* predicates a property over its argument *the book*.

(19) English (Indo-European)
   a. John is a teacher
   b. The book is red

With these definitions and cross-linguistic tendencies in place, the next subsection will draw attention to a number of languages which challenge these generalisations. In particular, a number of languages will be treated where the functions of reference and predication do not neatly map onto the lexical categories of nouns and verbs.

2.2.2 Challenges for the functional noun-verb distinction

As already shown for English in (19a-b), the functional noun-verb distinction as a distinction between lexemes used for reference and lexemes used for predication is not universally valid. In English, one could still argue that, since the noun *teacher* in (19a) needs the copula to be able to function as a predicate, this functional distinction still holds. A number of other languages, however, show this distinction to an even lesser degree. It has been claimed that languages as diverse as Nahuatl (Launey 1994, 2004), Straits Salish (Jelinek & Demers 1994), Tagalog (Lemaréchal 1989), Mundari (e.g. Hengeveld 1992), Enlhet (Kalisch 2009), and several Tupí-Guaranían languages (Queixalós 2006) all lack a syntactic basis for a noun-verb distinction.

In a language such as Nootka for example, both (20a) and (20b) are grammatical sentences. (20a) is the canonical mapping of (semantic) categories to syntactic functions that we would expect: the word denoting a time-stable entity (i.e. a typical semantic noun), *qo:ʔas* ‘man,’ functions as a referring expression (referring to a specific instance of a man in the extralinguistic world), and is an argument of the predicate *mamo:k* ‘work,’ which denotes an event.

(20) Nootka (Wakashan, Swadesh 1939: 78 via Queixalós 2006: 251)
   a. *mamo:k-ma* *qo:ʔas-ʔi*
      work-3SG.IND man-DEF
      ‘The man works’
   b. *qo:ʔas-ma* *mamo:k-ʔi*
      man-3SG.IND work-DEF
      ‘The worker is a man’

When, however, as in (20b), the lexemes are inserted in the syntactic frame the other way around, both of them take on a different function (Queixalós 2006: 251). The 'noun' *qo:ʔas*, when occurring in the first position of the clause, functions as a predicate ('to be a man'), whereas the 'verb' *mamo:k*, when following this predicate, functions as its argument (i.e. a referring expression, 'the one who works'). From this, one can conclude that there are few if any functional differences in this language between what we would prototypically call nouns and verbs. Both can function as arguments and both can function as predicates, without needing any morphological or syntactic increments (such as, for example, a noun needing a copula in order to function as a predicate, or a verb needing nominalising morphology in order to function as an argument). Languages of this kind have been called *monocategorial* (Evans & Osada 2005) or *flexible* (Hengeveld et al. 2004).

In Enlhet as well, according to Kalisch (2009: 122), nouns can function as predicates without needing any additional operations. In (21a), the noun *sekaʔa* 'child' functions as a predicate, even though it occurs in what seems to be a typical argument position. Nevertheless, the reverse does not hold: verbs need a nominalising suffix (such as the infinitive, usually *-ma* or *-ejʔ*) in order to function as arguments (21b). Therefore, all Enlhet lexemes can function as predicates: it is omnipredicative (in Launey's 1994 terms), but not monocategorial (since verbs cannot function as arguments, and thus still constitute a separate word class).

(21) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 122)

a. *aŋ-law-kek* *ła:* *n-sekaʔa*
F-get_angry-FACT.PRIM HOD F-child
'The child got angry. (lit.: She got angry. She is a child.)'

b. *wet-ʔa:* *nłoʔ* *ap-jew-ejʔ*
see-FACT.PRIM INT M-become_large-INF
'Did you see the/a big one?'

Functional equivalencies between nouns and verbs should not be taken to mean, however, that there is no basis at all in these languages for distinguishing between a category of nouns and one of verbs: they might very well be distinct on morphological grounds. Haude (2009: 331-2) argues, for example, that Movima nouns can take reduplication whereas verbs cannot. Therefore, even though they belong to one functional class, they should still be analysed as pertaining to different formal classes.

I will, in this paper, adhere to the criteria laid out in Evans & Osada (2005: 365-6), Hengeveld et al. (2004) and Rijkhoff & Van Lier (2013) in order to determine whether a language can be called monocategorial, omnipredicative, or flexible. The first criterion of Evans & Osada (2005: 365) is that of equivalent combinatorics: both 'verby' and 'nouny' words must be used with roughly equal frequency in predicative and referential function, without any additional morphological or semantic increments. If this holds in the two directions, a language can be called monocategorial. A less strict form of this criterion is adopted by Rijkhoff & Van Lier (2013). In their view, if all lexemes can grammatically occur as predicates and arguments (without necessarily doing so in the same frequency), this language has a flexible class of lexemes. If this requirement is only met unidirectionally, as in Enlhet, the language can still be called omnipredicative (since all lexemes can function as predicates), but no longer monocategorial.

Evans and Osada's (2005) second criterion states that the meaning of a lexeme used in a non-prototypical syntactic function should be straightforwardly derivable from the semantics of the lexeme plus the abstract semantics of the function. This holds true for the Nootka example in (20): the lexeme with the meaning 'work' means 'to perform the action of working' when it functions as a predicate, and 'be

one who performs the action of working' when it functions as an argument. For other alledgedly monocategorial languages, this is not always the case. In Mundari, for example, Evans & Osada (2005: 371) argue that the semantic leap between the referential and predicative use of the lexeme *hoRo* is too large. When used as an argument, it means 'a Munda person,' but when it is used as a predicate without any morphological or semantic alteration (22a), it means 'to speak the Munda language.'

(22) Mundari (Munda, Evans & Osada 2005: 371)

a. *ne dasi hoRo-a=eq*
this servant Munda-IND=3SG
'This servant speaks Munda'

b. *ne dasi hoRo tan-iq*
this servant Munda COP-3SG.S
'This servant is a Munda'

The only expected (and allowed) semantic modification that one could expect to result from the addition of the semantics of the predicative position is one from 'a Munda person' to 'to be a Munda person.' In order to receive this reading, however, the lexeme must obligatorily co-occur with the copula, as in (22b). Therefore, Mundari fails the criterion of semantic compositionality, and can in Evans & Osada's (2005) framework not be seen as a fully monocategorial language. Hengeveld et al. (2004) and Rijkhoff & Van Lier (2013) have a different view on this issue, however. Instead of seeing lexemes as having a specific sense which is then supplemented by the meaning of a syntactic slot, they see lexemes as having a vague sense existing of various components. Instead of adding to this vague sense the meaning of the syntactic slot in which the lexeme occurs, this syntactic slot highlights a number of these components and downplays others. In such an analysis of (22), it can be said that the predicate slot simply *highlights* the characteristic of speaking the Munda language, and *downplays* the characteristic of being born as a Munda person. Therefore semantic differences as those seen in the Munda example do not preclude an analysis of the language as possessing a functionally flexible lexeme class (Rijkhoff & Van Lier 2013: 14).

Thirdly, and lastly, Evans & Osada (2005: 378) argue that the functional equivalence of lexemes must be *exhaustive* in order for a language to be characterised as omnipredicative or monocategorial. In other words, the functional equivalence must be valid for *all* lexemes in a language, or at least the overwhelming majority of them, and not only a subset. It is, for example, not uncommon across languages that a number of kinship terms (prototypical nouns) can be used as predicates (Evans 2000). Even English can, to some extent, be said to show such a pattern: the words for *mother* and *father* can be used both as referents and as predicates, albeit with some more semantic leeway than Evans & Osada's (2005) semantic compositionality allows. Even if the semantic compositionality criterion held true for these words, however, English could still not be said to be an omnipredicative or monocategorial language, since other prototypical nouns (even those from the same semantic category, such as *sister*, *son* or *uncle*) cannot be used predicatively in this way.

All in all, based on these criteria it is possible to classify languages which show a fuzzy functional border between nouns and verbs along a continuum of monocategoriality. On the one end would be languages which show a functional equivalence between semantically prototypical nouns and verbs only in a subset of their lexicon, or languages where the semantic leap between the use of a lexeme as an argument and that as a predicate is too large. On the other end would be languages where *all* lexemes are distributionally equivalent, and the criterion of semantic compositionality holds: the true monocategorial languages. Somewhere in between, one could locate the omnipredicative languages: here, the criterion of distributional equivalence is only partially fulfilled, since nouny lexemes can function as

predicates, but verby lexemes cannot necessarily function as arguments. Such a continuum could then look as follows, with functional equivalence between noun and verb categories increasing from left to right:

partly fuzzy (e.g. not exhaustive through lexicon) – omnipredicative – monocategorial

2.3 The wave model of language diversification

As mentioned in section 1.3 above, one of the objectives of this thesis is to present insights into the historical diversification of the Enlhet-Enenlhet language family, and present additional evidence supporting or weakening certain elements of the internal structure of the language family proposed by Unruh & Kalisch (2003). Ever since the 19th century German Neogrammarians, the dominant model of representing genetic relations between languages has been the family tree model, resulting in a visual representation of these genetic relations as in figures 1 and 2 in subsection 1.2.1.

Such a model, however, can be argued to contain a number of substantial flaws. It assumes that the main cause of language diversification is the sudden division of populations speaking one protolanguage, after which these languages start innovating in their own separate ways (François 2014: 132). What it does not take into account, according to François (2014), is situations where a protolanguage gradually diversifies into a dialect continuum where speakers of the separate languages stay in contact with each other. This is, according to Unruh & Kalisch (2003), the situation which most closely resembles the difersification of the Enlhet-Enenlhet languages. In this section, I will summarise François' (2014) argument against the use of a family tree model in such situations, and explain the alternative model he proposes and which will be applied here: the wave model of language diversification.

When presenting a family tree-based diversification of the Enlhet-Enenlhet language family (such as the one proposed in figure 2), one makes a number of assumptions. For example, one assumes that at a certain point in time, there was a community of speakers with a shared language: Proto-Enlhet-Enenlhet. Furthermore, one assumes that, at a certain point in time, a population split took place, resulting in two speaker communities. The languages of these communities then, over the years, adopted a number of innovations, resulting in a Proto-Western-Enlhet-Enenlhet language (the result of Proto-Enlhet-Enenlhet and innovations A and B, for example), and a Proto-Eastern-Enlhet-Enenlhet language (resulting from Proto-Enlhet-Enenlhet undergoing innovations C and D). Later, the same process repeated itself, resulting in, for example, the Enlhet language as the result of Proto-Western-Enlhet-Enenlhet and innovations E and F; and the Enxet language as the result of this protolanguage and innovations G and H.

A vital point now, of such an argumentation, is that all innovations are necessarily limited to one (sub)branch of the family (François 2014: 164). If innovations A and B are used to posit a Western branch of the family (which possesses these innovations) and an Eastern branch (whose languages do not possess them), it should not be the case that any language of the Eastern branch shows one of these innovations. Unfortunately, this is not the case in the Enlhet-Enenlhet language family. As claimed by Unruh & Kalisch (2003) and further evidenced in chapter 6 of this thesis, there are a number of innovations which are shared by languages of the Eastern branch and of the Western branch (for example, features shared by Enlhet, Guaná and Enenlhet, but not by Sanapaná). Such features, then, pose a challenge for the family tree model.

If such shared features between languages of different branches of a language family are found, a proponent of the tree model can resort to three explanations (François 2014: 166): he can present the feature as a shared retention in these languages, which was lost in the other sister languages (for example, the feature was maintained in Enlhet, Guaná and Enenlhet, but lost in Enxet, Sanapaná and Angaité); he can present the feature as an independent innovation in the languages of different sister branches (which is not a very economical solution); or he can present it as a contact-induced similarity that arose after the languages in question had already fully diversified.

The tree model, however, does not take into account situations where features spread between different languages of the same family *while they were in the process of diverging*, i.e. when they did not as yet exist as independent languages. In other words, it could have been the case that the innovation shared by Enlhet, Guaná and Enenlhet originated while these present-day languages were still mutually understandable idiolects of one single language, starting in one part of the speaker community (say amongst the Guaná-idiolect speakers), then spreading amongst the Enlhet-idiolect speakers and Enenlhet-idiolect speakers, but not reaching the speakers of the Enxet, Sanapaná and Angaité idiolects. In such a situation, the present-day Enlhet, Guaná and Enenlhet languages must be posited to form together a valid genetic subgroup of the Enlhet-Enenlhet language family, an observation which cannot be reconciled with the family tree proposed in figure 2 above. To use François' (2014: 167) words:

> While it may be the case that dialects KL [for example, Enlhet and Enxet] have shared together one set of innovations and MNO [for example, Guaná, Enenlhet, Sanapaná and Angaité] another one, it is perfectly plausible that dialects L and M [for example, Enlhet and Enenlhet] could also undergo their own set of shared innovations, during the same historical period. This is how the process of language-internal diffusion, the ultimate source of genealogical relations in languages […], can give birth to subgroups that crosscut each other: KL; LM; MNO… Such a dialect-chain situation, and more generally dialect continua and linkages […], form the Achilles' heel of the Tree Model.

If the present-day Enlhet-Enenlhet languages, then, resulted from a set of mutually intelligible idiolects which stayed in contact throughout their diversification process and consequently created a web of intersecting innovations, constituting different genetic subgroupings within the language family, a different representational model should be adopted. Under the wave model of language diversification, first proposed by Schmidt (1872), linguistic innovations are no longer seen as happening in one intermediate node of a language family tree and then being passed on to its descendents without spreading into any other languages. Instead, they are seen as 'waves' starting in a central point (a certain community of speakers of an idiolect), and then spreading out into neighbouring communities.

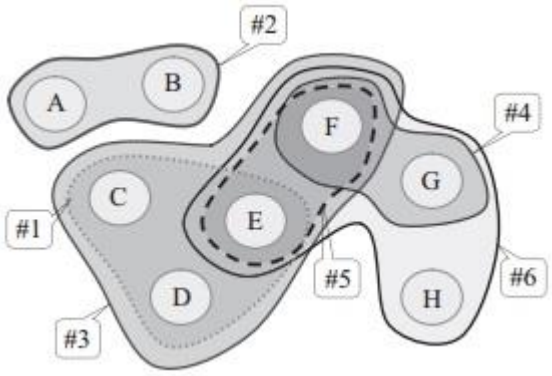

Figure 4: Innovations spreading through a hypothetical dialect continuum (copied from François 2014: 169)

This can be illustrated by means of figure 4, representing the geographical distribution of eight hypothetical languages (A, B, C, D, E, F, G and H) which descended from a hypothetical protolanguage. These languages, or more accurately, the idiolects of Proto-ABCDEFGH from which they evolved, took part in six hypothetical innovations, represented by the isoglosses grouping different languages together. Innovation 2, for one, only took place in idiolects A and B. This scenario can be represented accurately in a family tree: Proto-ABCDEFGH would then have two branches, proto-AB and proto-CDEFGH, and every language would only pertain to one of the subbranches.

The other innovations do not allow for such clean, binary branching scenarios, however. Innovation 1 would prompt the linguist to establish a proto-CDE branch (which takes part in it) versus a proto-FGH branch (the members of which do not possess it). Innovation 6, on the other hand, would force a proto-EFGH versus proto-CD interpretation. The main question here is, to which branch of the language family does E belong? It shares features exclusively both with C and D on the one hand, and with F, G and H on the other. Yet, a tree model does not allow us to classify E in both these branches at the same time.

In the wave model, however, it is perfectly acceptable for a language to belong to different subgroups, or in other words, subgroups of language families are allowed to (partially) overlap (François 2014: 170). Specifically, every set of two or more languages of a language family which share at least one innovation is in this model seen as a genetic subgroup which had throughout history a certain degree of shared development. As such, language E in figure 4 would be said to belong to subgroup CDE, subgroup EF, and subgroup EFGH at the same time. In the Enlhet-Enenlhet scenario posited above, it could then for example be the case that Enlhet belongs to an Enlhet-Enxet subgroup based on a number of features shared by these languages but absent in Guaná, Enenlhet, Sanapaná and Angaité; but this same language could belong to an Enlhet-Enenlhet-Guaná subgroup based on a a different set of innovations.

Of course, not all genetic subgroupings of a language family are equally strong. In figure 4, for example, languages E and H on the one hand, and E and F on the other form equally valid genetic subgroups: the former two participate in innovation 6, whereas the latter participate in innovations 3, 5, and 6. Nevertheless, the fact that E and F share three features whereas E and H only share one reflects a historical scenario where the speakers of E and F were in closer or more extended contact than those of E and H, a difference which, ideally, one should be able to quantify (François 2014: 174). To that end, François (2014: 180) proposes the method of Historical Glottometry, where the measures of *cohesiveness* and *subgroupiness* are used in order to establish which genetic subgroups of a language family are the strongest.

The cohesiveness of a subgroup is established by calculating "the proportion of evidence supporting that subgroup with respect to the entire set of relevant evidence" (François 2014: 180). Henceforth, I use p as symbolising the number of supporting innovations for a certain subgroup (i.e. the number of innovations that occur in all the members of this subgroup), and q as symbolising the number of conflicting innovations for this subgroup (i.e. the number of innovations that occurs in one or more, but not all, of the members of this subgroup, and one or more languages outside this subgroup). In order to calculate the cohesiveness c of a certain subgroup G, then, one must apply the formula $c_G = p / (p + q)$.

As a concrete example, I will now calculate the cohesiveness of the subgroups [E, F] and [E, H] of the hypothetical language family of figure 4. For [E, F], $p = 3$: the two languages share innovations 3, 5 and 6. The q for this subgroup is 2: innovation 1 involves C, D, and E but not F, and innovation 4 involves F and G but not E. Both these innovations weaken the [E, F] subgroup-hypothesis. When applying the formula for cohesiveness, it can be seen that $c_{[E, F]} = 3 / (3+2) = 0{,}6$. For [E, H], next, $p = 1$ (supported by innovation 6) and $q = 3$ (weakened by innovations 1, 3 and 5), such that $c_{[E, H]} = 1 / (1+3) = 0{,}25$.

These numbers should be interpreted as follows (François 2014: 181): whenever an innovation affected either language E or language F, there is a 60% chance that it affected the other one as well. On the other hand, whenever an innovation affected either language E or language H, the chance that it affected the other one is only 25%. As such, it can be hypothesised that the speakers of the ancestor of language E had closer linguistic (and by extension social) ties to those of the ancestor of language F than to those of the ancestor of language H.

The second measure introduced by François (2014: 181), is the eventual indicator of the overall strength of a subgroup. This measure, called subgroupiness (*s*), is calculated by multiplying a subgroup's cohesiveness by the number of times *e* this subgroup shares an exclusive innovation (i.e. one not shared by any of the other family members): $s_G = e * c$. As such, the [E, F] subgroup of the hypothetical language family portrayed above presents a subgroupiness rate of 1 * 0,6 = 0,6. Now imagine that there had been an additional isogloss on figure 4, representing an innovation 7 which only affected languages E and F. In that case, the *c* of [E, F] would have been 4 / (4+2) = 0,66; and its *s* would have been 2 * 0,66 = 1,32.

In this way, one can bring all the available evidence to bear on how strong a subgroup is: one takes into account both the absolute number of times this subgroup is attested, and the relative measure of how often this subgroup is supported divided by how often it is contradicted. Therefore, the measure of subgroupiness is the one that will be used in chapter 6 of this thesis to determine which subgroups of the Enlhet-Enenlhet language family are supported by the strongest evidence, and whether or not these observations corroborate Unruh & Kalisch' (2003) posited Western, Northeastern, and Southeastern Enlhet-Enenlhet subgroups.

Chapter 3 A Brief Introduction to Sanapaná Grammar

In this chapter, I present some aspects of Sanapaná grammar that will become relevant in chapter 4 for the discussion of the particles *łke, ałta* and *hata*. I start with a brief typological overview of some general features of Sanapaná, followed by some notes on the phonology and orthography. Aferwards, I present a survey of nominal and verbal morphology inasmuch that this pertains to the discussion in chapter 4. The hypotheses presented about Sanapaná in this chapter are rather heavily inspired by Kalisch' (2009), Unruh & Kalisch' (1999), and Unruh et al.'s (2003) work on the other Enlhet-Enenlhet languages. They are therefore necessarily rudimentary, and await confirmation or contradiction by further data collection. In the mean time, I take full responsibility for any errors in my analysis.

3.1 Typological characteristics

I start my introduction to Sanapaná grammar by briefly surveying a small, typologically inspired, set of general features of the language: constituent order, argument marking, and the delimitation of word classes. Constituent order, firstly, is quite variable. When both arguments of a transitive clause are overtly expressed through NPs, the most unmarked order seems to be VOS: this order can be seen in (23a). This constituent order is quite rare in natural discourse, however. In many instances, verbal arguments can be suppressed, because they are identifiable from the person marking morphology on the verb, and from the context (23b).

(23) Sanapaná (FN3.3; 3.1)

a. *apk-apas-ke=łta* *ap-ketka* *en-japoŋ neteŋ*
2/3M-send-REAL=PHOD 2/3M-child PL-father top
'Our father above sent his son.'

b. *vakka* *ał-to-ma-kha=łta* *ap-te-j-akmok*
cow.SP[2] PL-eat-INF-LOC=PHOD 2/3M-be_born-STX-V2.IPF
'You were born in the food trough of the cows.'

[2] If the tress in Spanish loans is maintained and results therefore in an abnormal position for Sanapaná (i.e. preceding the last syllable), this is indicated by means of an acute accent on the stressed syllable. In this case, stress is adapted to the Sanapaná system (*vak'ka*).

In addition, it is not at all rare that the verb does not occupy the first clausal position. Both arguments (24a) and adjuncts (24b) can precede the verb without influencing the grammaticality of the utterance.

(24) Sanapaná (FN2.88; 2.69)

a. *tenjo* *koʔo* *na* *as-tejan-me-ʔaw=łta* *metaʔa* *naʔa*
then PRO:1SG PRES 1SG-look-INF-INTNS=PHOD always PRES

as-ne-k-a-maha *konne* *peskeska*
1SG-sit-STX-IPFV-EXT bottom shade
'Then I was, as always, looking [at the football game] and sitting in the shade.'

b. *kełwoje=łta* *konj-eheʔ* *eŋ-wet-aʔ* *eneŋ-koʔo*
moment=PHOD run-REAL.V1 1PL.O-see-REAL.V2 PL-PRO:1SG
'Immediately, it [the ostrich] ran away when it saw us.'

Consequently, it can reasonably be assumed that constituent order obeys pragmatic criteria rather than grammatical ones: there are hardly any grammatical restrictions on constituent order (except for a slight preference for verb-initial orders), but focussed constituents are often fronted to occur in clause-initial position.

Sanapaná does thus not make use of a fixed constituent order to disambiguate A and O arguments in transitive clauses. Neither does nominal morphology provide any help in this task: as shown in (23a), nouns take prefixal morphology for gender and number, but do not have a case-marking slot. Examples (25a-b) show this more explicitly: neither the pronoun *koʔo* nor the noun *semhen* receives any case-marking morphology to indicate their function as an A or an O. The only nominal morphology present is the (in rapid speech optional) gender prefix *a-* on the noun *semhen*, indicating its feminine gender.

(25) Sanapaná (FN1.61)

a. *a-semhen* *e-tjene-ʔaw* *(koʔo)*
2/3F-dog 1SG.O-look-INTNS[3] PRO:1SG
'The dog is watching me.'

b. *(koʔo)* *as-tjene-ʔaw* *semhen*
PRO:1SG 1SG-look-INTNS dog
'I am watching the dog.'

The way in which Sanapaná disambiguates the A and O arguments of transitive clauses, then, is by means of cross-referencing prefixes on the verb. This can be seen throughout examples (23-25): in (25a-b), for example, the *e-* and *as-* prefixes both refer to a first person argument, but the former indicates that it functions as an O whereas the latter indicates that it functions as an A. This same *as* prefix is found on an intransitive verb in (24a), indicating that these markers show nominative-accusative alignment: the same morphemes are used for S and A arguments, as opposed to O arguments which are marked differently. An extra complication for this system is posed by the person hierarchy that affects it, a topic that will be treated in some more detail in subsection 3.4.1.

The last issue treated in this section is that of Sanapaná word classes. Sanapaná can be argued to possess only two open classes of lexemes: one of verbs and one of nouns. There are many significant (mainly morphological) differences between these two classes: the former are marked for TAM (such

[3] The *intensive* suffix, according to Unruh et al.'s (2003) analysis of Enenlhet, can amongst others function as a typical intensifier, expressing that the action expressed by the verb took place with increased intensity; it can indicate that an action took place unintentionally; and it is lexicalised with certain verbs. Since I do not have any forms of the verb *tjene* without this suffix, I assume, awaiting additional data, that the latter analysis applies here.

as the realis suffix *-ke* in 23a), location and direction, and subordination (such as the V2 morpheme *-aʔ* in 24b), whereas the latter are not. Nevertheless, there are also a number of similarities: both classes make use of the same prefixes for argument marking (marking the person and gender of the possessor or the referent of a noun, and of the S, A and O argument of a verb), for example. Sanapaná nominal and verbal morphology will be treated in some more detail in sections 3.3 and 3.4.

Many notions that are typically expressed through members of other word classes in other languages can in Sanapaná be seen to pertain to one of these two open classes. Most prototypically adjectival concepts are expressed through nouns (e.g. *ak-wanjam* 'oldness,' *as-wanjam* 'my oldness' i.e. 'I am old.'), as are most concepts that would in English be expressed through prepositions (such as *neteng* 'top, sky,' in 23a above, which is also used to express the locative relation 'on top of'). Many adverbial concepts are nominal in Sanapaná as well, such as the word for 'yesterday' *aknem=ałta* (lit. sun/day-PHOD), and the word *nen-ta:men* (IMPERS-past), which can be used both with the meaning 'before' (i.e. preposition-like) or meaning 'in the past' (i.e. adverbial-like).

Word classes that do exist and have a fair number of members in Sanapaná are conjunctions (*kanhan* 'and,' *tenjo* 'then'), demonstratives (*heŋkaʔe* 'this, here,' *namʔa* 'that, over there') and particles. Three members of this last category, which we saw represented in the example sentences of this chapter by amongst others *anak*/*naʔak* (presentative), will be the focus of the next chapter.

3.2 Phonology and orthography

Sanapaná makes use of a relatively limited phoneme inventory. According to Gomes (2013: 82; 105), the language has at its disposal 3 vowel phonemes, and 13 consonant phonemes. The vowel phonemes are a low central vowel /a/, a mid front vowel /e/, and a mid back vowel /o/, all exemplified in the word *enjapoŋ* 'our father' in (23a) above (table 2). The two mid vowels can be raised in certain phonetic contexts. For example, when *neteŋ* is followed by the presentative particle *ana(ʔak)*, this results in the phonetical sequence [*netiana*]. I encountered, however, no minimal pairs contrasting *e* with a different front vowel or *o* with a different back vowel.

	Back	Central	Front
Close			
Mid	o		e
Open		a	

Table 2: Sanapaná vowel inventory

Vowel length seems to play a role in Sanapaná as well, especially relating to number, as noted by Gomes (2013: 107) for the plural of the diminutive. Several nouns, especially those referring to body parts and family members, take a *-ʔV-* infix in their last syllable for pluralisation: *ansetkok* 'child' – *ansetkoʔok* 'children' (see also section 3.3). Additionally in verb forms such as *maʔe etok* 'you (M.SG)/he will eat' – *maʔe ełtoʔok* 'you/they (M.PL) will eat,' the same phenomenon, insertion of a vowel *-o-* in the last syllable is one of the ways in which the plural is marked. Since intervocalic glottal stops are often lost in rapid speech, such forms such as the last one in this list may surface as *ansetko:k* and *ełto:k*, resulting in a vowel length distinction marking number.

Furthermore, several lexemes (such as *pesa:sep* 'night') are consistently produced with a lengthened vowel, and a learner is corrected when he unobservantly pronounces them with a short vowel. Others, such as *kelasma* 'fish,' consistently receive a short vowel. Since I do not have any minimal pairs of this kind distinguished by vowel length, I will not make any claims here about their phonetic or phonemic status, but rather leave the discussion of Sanapaná vowel length for future investigation.

The consonant phonemes posited by Gomes (2013) can be divided into five series. There are four plosives (bilabial /*p*/, alveolar /*t*/, velar /*k*/ and glottal /*ʔ*/), three nasals (bilabial /*m*/, alveolar /*n*/ and palatal /*ɲ*/), three fricatives (alveolar /*s*/, velar /*ɬ*/ and glottal /*h*/), and three approximants (alveolar lateral /*l*/, palatal /*j*/ and labiodental /*w*/).

Although I did not focus on phonology during my own fieldwork, I do see a number of issues with this classification. First and foremost, Gomes' (2013) classification of the sound he transcribes as <hl> (as in <pehlten> 'month') as a velar fricative, and of the /*w*/ as as a labiodental approximant strike me as odd. The first one, to my ear, and according to Kalisch' (2009: 110) comments on Enlhet phonology, is better characterised as a voiceless alveolar lateral fricative, whereas the second one seems to be labiovelar rather than labiodental.

Secondly, with regards to his nasal phonemes, I would posit a velar nasal phoneme /*ŋ*/, rather than seeing it as an allophone of /*n*/ before a velar phoneme. My reason for this is that both sounds seem to occur preceding the labiovelar approximant /*w*/ (a context in which one would expect only the velar nasal if this was a conditioned allomorph), and word-finally: *aŋwajka* 'cooking pot' – *enwałwok* 'stomach,' *nenta:men* 'past' – *enjapoŋ* 'our father'. These examples seem to suggest that the velar and alveolar nasals are not in complementary distribution, and they are, additionally, not in free variation. Therefore I posit /*ŋ*/ as a phoneme of Sanapaná.

Furthermore, I am not entirely convinced of the classification of the palatal nasal /*ɲ*/ as a phoneme, rather than as a sequence of /*n*/ followed by a glide /*j*/. Since Gomes (2013) does not provide argumentation for this analysis, and I do not have conclusive evidence yet in my own data, I leave the jury out on this phoneme. For now, however, it seems more economical to not have to posit an extra phoneme and analyse this sound as a /*nj*/ cluster. My phonemic chart of Sanapaná is represented in table 3.

	Bilabial	Alveolar	Palatal	Velar	Labiovelar	Glottal
Stops	p	t		k		ʔ
Nasals	m	n		ŋ		
Fricatives		s				h
Lateral fricative		ɬ				
Approximants		l	j		w	

Table 3: Sanapaná consonant inventory

The orthography typically used in La Esperanza to write the language down was developed in the latter half of the 20th century by the Asociación de Servicios de Cooperación Indígena-Mennonita (ASCIM), an organisation with which the community of La Esperanza is in close contact with regards to, amongst others, the management of their lands (Gomes 2009a: 374, personal observation). The most important features of this Spanish-inspired orthography are the following (Gomes 2009a: 376; personal observation):

i. The vowels and most consonants are represented by the same grapheme that is used in Spanish: <a, e, o, p, t, k, m, n, s, h, l>
ii. The glottal stop /*ʔ*/ is represented by <'>
iii. The velar nasal /*ŋ*/and the lateral fricative /*ɬ*/are represented by digraphs <ng> and <hl>

iv. The semivowels /*j*/ and /*w*/ are represented by <y, i> and <v, u>
v. The sequence /*kh*/ is often represented as <h>
vi. Lengthened phonemes are represented by double graphemes

Despite the existence of a more or less standardised Sanapaná ortography, all examples in this thesis are rendered in IPA, using the symbols found in tables 2-3 above.

3.3 Nominal morphology

Sanapaná nominal morphology is significantly less complex than that of the verb. Often, nouns are used simply as a stem, without any additional morphology at all. This can be seen in (26a). The noun *mesa*, a loan from Spanish, is used without any additional morphology, as is the noun *neteŋ*. This example also shows that Sanapaná does not have articles at its disposal (although the use of the numeral *łema* 'one' as an indefinite article is on the rise, possibly under influence of Spanish).

(26) Sanapaná (FN2: 116)

a. *ja:met-atkok aʔ-jetn-ema neteŋ mésa*
branch-DIM 2/3-lie-INF top table.SP
'The small branch is lying on top of the table.'

b. *as-tawa*
1SG-spouse
'my spouse'

c. *a-semhen*
F-dog
'the/a dog'

Example (26a) also shows one of the ways in which nouns can take suffixes: the suffix *-atkok*, plural *-atkoʔok* is a diminutive. The addition of a vowel infix (often *-o-*) in the last syllable of the noun is the only strategy Sanapaná has at its disposal for overt nominal pluralisation. In addition to nouns modified with the diminutive suffix, other nouns pluralising in this way are *ahaʔtek* 'eye' – *ahaʔtaok* 'eyes,' *ahajkok* 'ear' – *ahajkaok* 'ears,' *emek* 'hand' – *emeok* 'hands.' Nouns can also take a suffix *-(ŋk)oje*, which arguably functions as an intensifier (as in *sosokh-oje* 'tomorrow,' from *sokha* 'moment' and *mokhoje-ŋkoje* 'very far' from *mokhoje* 'distance').

Furthermore, nouns can be marked with personal prefixes (the same ones as are used for verbs). With certain nouns (for example, body part terms and certain family members), the use of a personal prefix referring to the possessor of the noun is obligatory (26b), forming thus a class of inalienably possessed nouns. Other nouns can optionally take a prefix to make the grammatical gender of the noun explicit (26c). A last comment that must be made about nouns is that they can be used as predicates without undergoing any morphological or syntactic operations: the nouns in (26b-c) can equally well constitute a full utterance, meaning 'she is my spouse' and 'it is a dog'.

3.4 Verbal morphology

Just as the other Enlhet-Enenlhet languages such as Enenlhet (Unruh et al. 2003) and Guaná (Unruh & Kalisch 1999), Sanapaná makes extensive use of verbal suffixing morphology to mark categories of TAM, location and direction, subordination, and valency-changing operations such as causativisation. In combination with the multitude of verb classes which make use of different morphological material to mark these categories, this makes Sanapaná verbal morphology extremely complex. Unfortunately, due to time constraints, the data collected during my two-month field trip to La Esperanza are too limited to allow a full analysis of this morphological system. Therefore, I limit myself in this section to the discussion of verbal cross-referencing and TAM, in the verb forms that are used when the main verb of a clause occurs in initial position and is not subordinated to another predicate (called *primative* by Unruh & Kalisch 2003 and Kalisch 2009, I will here label it the V1-form of the verb). The description of the V2-forms and subjunctive forms (both modes of subordination), and the numerous locative and directional categories is left for later work.

I start my discussion of Sanapaná verbal morphology with a survey of the cross-referencing prefixes used to indicate the arguments of the verb, and then move on to the three TAM categories that are relevant for the present study.

3.4.1 Verbal cross-referencing

Sanapaná verbs, when used in a syntactic context, obligatorily take a person marking prefix. In most cases, this prefix refers to the S or A argument. The most common set of prefixes is given in table 4 (although other prefixes occur for different verb classes, which have not been systematised yet) and demonstrated in (27a-f) for an intransitive verb.

(27) Sanapaná (FN2.19)

a. *as-taw-ke*
1SG-eat-REAL.V1
'I am eating'

b. *ap-taw-ke*
2/3M-eat-REAL.V1
'you (M.SG) are/he is eating'

c. *an-taw-ke*
2/3F-eat-REAL.V1
'you (F.SG) are/she is eating'

d. *eɬ-taw-ke*
PL-eat-REAL.V1
'we are eating'

e. *apk-eɬ-taw-ke*
2/3M-PL-eat-REAL.V1
'you (M.PL)/they (M) are eating'

f. *aŋk-eɬ-taw-ke*
2/3F-PL-eat-REAL.V1
'you (F.PL)/they (F) are eating'

As argued by Gomes (2013: 284) and exemplified in the forms above, Sanapaná verbal cross-referencing makes a distinction between first person and non-first person arguments. First person singular arguments are, on the verb *taw*[4] 'to eat,' marked with a prefix *as-*. For non-first person arguments, a distinction is made based on gender rather than number: second or third person masculine takes an *ap-* prefix, as opposed to an *an-* prefix for second and third person feminine. Non-singular first person, non-first person masculine and non-first person feminine all take a prefix *eɬ-* before the verb stem. This is preceded, for the latter two, by the non-first person singular prefix and an added *k-* segment which appears between the person prefix and a following vowel. Since there is only one non-singular number in Sanapaná, I will henceforth refer to it as 'plural.'

	Realis	Irrealis
1SG	*as-/e-*	*o-*
2/3SG.M	*ap-*	*e-*
2/3SG.F	*an-*	*aŋko-*
(1)PL	*e(ɬ)-*	*o-ɬ-*
2/3PL.M	*apk-eɬ-*	*e-ɬ-*
2/3PL.F	*aŋk-eɬ-*	*aŋko-ɬ-*

Table 4: Most common personal prefixes in Sanapaná

In monotransitive clauses, secondly, verbs most of the time cross-reference their A argument by means of a prefix from the same set that is used to mark S arguments (28a-c).

(28) Sanapaná (FN1.61)

a. *koʔo* *as-tjene-ʔaw* *semhen*
PRO:1SG 1SG-look-INTNS dog
'I am watching the dog.'

b. *koʔo* *as-tjene-ʔaw* *łejap*
PRO:1SG 1SG-look-INTNS PRO:2/3SG.M
'I am watching you (M.SG).'

c. *łejap* *ap-tjene-ʔaw* *semhen*
PRO:2/3SG.M 2/3M-look-INTNS dog(F)
'You (M.SG) are/he is watching the dog.'

Nevertheless, when the O argument is a first person, it is usually this O argument that is indicated with a prefix instead of the A (29a-d). I analyse this as a case of hierarchical alignment. According to Siewierska (1998: 9), languages showing hierarchical alignment make use of a person hierarchy, and whichever argument ranks highest on this hierarchy "receives special treatment, the details of which vary from language to language."

(29) Sanapaná (FN1.61; 3.6; 1.119)

a. *a-semhen* *e-tjene-ʔaw*
2/3F-dog 1SG.O-look-INTNS
'The dog is watching me.'

b. *łejap* *e-tjene-ʔaw*
PRO:2/3SG.M 1SG.O-look-INTNS
'You (M.SG) are/he is watching me.'

[4] The citation form used throughout this thesis to refer to a verbal lexical item is the verb stem used in the present tense paradigm.

c. *e-jamela-jam* *jesus* *ak-weske*
1SG.O-love-INF Jesus 2/3F-chief
'The Lord Jesus loves me.'

d. *e-taɬ-e* *a-semhen*
1SG.O-bite-REAL.V1 2/3F.dog
'The dog bit me.'

The Sanapaná data in (28-9) could be explained accordingly. One can posit a first person > non-first person hierarchy, and a condition which states that there is one prefixal slot for person marking on the verb, which is filled by the highest-ranking argument. In this way, whenever a first person is present in the interaction, acting upon or being acted upon by a non-first person, this first person is prefixed to the verb stem. When the first person occurs in A function, the *as-* prefix is used (28a-b), as opposed to the *e-* prefix when it functions as an O (29a-d). When no first person is present in the interaction, both arguments are evenly ranked on the Sanapaná person hierarchy, and it is always the A that is cross-referenced: in (28c), for example, the masculine singular prefix is used rather than the feminine one, indicating that the masculine singular pronoun *łɛjap* is the subject of *aptjeneʔaw* rather than the feminine noun *semhen*.

3.4.2 Verbal TAM

In terms of morphological manifestations of TAM on the verb, one must bear in mind the intricate correlation of tense, aspect and modality in Sanapaná, where especially tense and mood are closely intertwined. Three morphological categories can be distinguished that appear to convey temporal information. The category which is used most often with future time reference in Sanapaná is formed simultaneously by means of an auxiliary *mVʔ(V)* which precedes the verb and makes a distinction between first person and non-first person, a personal prefix on the verb stem distinct from the set that we saw earlier, and a suffix added to the stem of the main verb. The last vowel from the auxiliary is typically elided when it comes in contact with the first vowel of the personal prefix. Both vowels are underlyingly present, however: (31b) below provides evidence for the prefixal vowel, whereas (43a) in section 4.1 provides evidence for the last vowel of the auxiliary.

(30) Sanapaná (FN2.19-33; 2.86)

a. *moʔo* *o-t-ok* *koʔo*
1SG.AUX.IRR 1-eat-IRR PRO:1SG
'I am going to eat.'

b. *maʔe* *e-t-ok* *anhan* *sosokha*
2/3.AUX.IRR 2/3M-eat-IRR CONJ tomorrow
'You (M.SG)/he will also eat tomorrow.'

c. *maʔa* *aŋko-t-ok* *anhan* *sosokha*
2/3.AUX.IRR 2/3F-eat-IRR CONJ tomorrow
'You (F.SG)/she will also eat tomorrow.'

d. *maʔa* *o-ɬ-t-oʔok*
1PL.AUX.IRR 1-PL-eat-IRR.PL
'We will eat.'

e. *maʔe* *e-ɬ-t-oʔok*
2/3.AUX.IRR 2/3M-PL-eat-IRR.PL
'You (M.PL)/they (M) will eat.'

f. *maʔa* *aŋko-ɬ-t-oʔok*
2/3.AUX.IRR 2/3F-PL-eat-IRR.PL
'You (F.PL)/they (F) will eat.'

As can be seen in (30a-f), both the auxiliary and the prefix distinguish first person, non-first person masculine and non-first person feminine. Neither the auxiliary nor the personal prefixes distinguish number: the plural[5] is indicated through an extra prefix *ɬ-* between the person prefix and the verb stem, and by replacing the single vowel in the suffix with two vowels separated by a glottal stop. The auxiliary, furthermore, is not obligatory. The forms in (31a-b), both found as translations of the Spanish *voy a dormir* 'I am going to sleep,' and which were judged to be equally grammatical, show this.

(31) Sanapaná (FN2.113)
a. *moʔo* *(o)-jt-ek*
1.AUX.IRR 1-sleep-IRR
'I am going to sleep.'

b. *o-ʧen-e=ta*
1-sleep-IRR=FUT
'I am going to sleep.'

Whereas in (31a), the auxiliary and the suffix together mark the TAM category, in (31b) the auxiliary is absent, and the person-marking prefix and the suffix are the only elements that convey TAM information. An additional clitic *=ta*[6] appears as well, which triggers a morphophonological process of deletion of the final *-k* of the *-ek* suffix. The alteration in the verb stem is due to a process of metathesis which is characteristic of a number of dialects of Sanapaná,[7] but has no significant effect on intelligibility.

Now that the morphological formation of this TAM category has been explained, I present a brief note on its semantic status. Gomes (2013: 311), for instance, argues that the basic meaning of this category is an aspectual one. In his analysis, the distinction found between the sentences in (30-1) and those in (27-9) is one between a prospective aspect, which refers to an event projected to happen at an indefinite time in the future; and a non-prospective aspect.[8]

I do not wholly agree with this analysis, however. I believe that what Gomes (2013) terms 'prospective aspect' can better be interpreted as a modal category. It is the case, namely, that the *mVʔV* auxiliary is also often used in utterances with an explicit deontic modal load. In (32a-b), for example, the auxiliary *moʔo* adds a clear meaning of desire to the semantics of the main verb.

[5] According to Kalisch (2009) and Unruh et al. (2003), the Enlhet-Enenlhet languages do not mark classical plural number on the verb. Instead, what I call the plural prefix here can have scope over either the A argument, the O argument, or the event as a whole. Therefore, the aforementioned authors adopt the term *distributive*. Since I have insufficient data to accurately describe the behaviour and semantics of this prefix in Sanapaná, I label it *plural* for now.

[6] This is a cliticised form of the particle *hata*, to be discussed in more depth in chapter 4.

[7] According to reports of native Sanapaná speakers in La Esperanza and John Stucky from New Tribes Mission (p.c. 2017), around five to six different variants of Sanapaná are spoken in the community. This is conceivably due to the fact that Sanapaná from different regions, and corresponding regional varieties of the language, were brought together in the La Esperanza mission in the 1960s. Consequently, lectal variation within the Sanapaná language would be a fascinating topic for further research.

[8] Terminology translated by me from the original Portuguese in Gomes (2013).

(32) Sanapaná (FN2.18)

a. *moʔo* *o-teʔpoŋ* *koʔo* *paga* *naʔak*
1.AUX.IRR 1-hit.IRR PRO:1SG mosquito PRES
'I want to kill a mosquito.'

b. *moʔo* *o-jase-ŋwom-ok* *nenɬet-ana* *a-pajoma*
1.AUX.IRR 1-learn-IBILOC[9]-IRR person-PRES 2/3F-language
'I want to learn Sanapaná.' (lit. 'I want to learn the language of the people.')

(33) Sanapaná (FN1.23)

a. *maʔe* *e-ɬn-ek*
2/3.AUX.IRR 2/3M-sit-IRR
'Are you (M.SG)/is he going to sit?/ Do you (M.SG)/does he want to sit?'

b. *moʔo* *o-ɬn-ek*
1.AUX.IRR 1-sit-IRR
'I want to sit/I am going to sit.'

Additionally, this auxiliary is often ambiguous between the future reading and the deontic reading discussed before. This is for example the case in (33a-b). Furthermore, in Enenlhet the verb *makʔak*, which is used in this context as an auxiliary, can also be used as a main verb with the meaning 'to want' (34). Therefore, it seems clear to me that the TAM category under discussion here presents a mix of modal and temporal semantics, which we saw in chapter two is not uncommon especially for categories used with future temporal reference. Since the auxiliary and suffix under discussion here can both give the verb an interpretation of 'desire' and one of 'unrealised event,' I analyse this Sanapaná category synchronically as irrealis, with the caveats that, when it is accompanied by temporal adverbials such as *sosokha* 'tomorrow' in (30b-c), a more strictly temporal reading is triggered.

(34) Enenlhet (Enlhet-Enenlhet, Unruh et al. 2003: 76)
kelasma *ɬna* *m-makʔak* *meme*
fish RPT 2/3F-want mother
'Mother said she wants fish.'

Of course, it is rather common in languages of the world for grammatical elements conveying modal categories to grammaticalise into tense markers. A few well-known examples of this process are the grammaticalisation of the Latin auxiliary *habere* (conveying a deontic modal load of obligation) into the future tense markers in the present-day Romance languages, or the English auxiliary *will*, originally conveying a modal meaning of desire, into a fully-fledged auxiliary only used for future tense marking (see amongst others Hopper 1991: 25-29). Whereas it is quite possible that the Sanapaná auxiliary under discussion here is undergoing a process of grammaticalisation and might in the future come to be used as a pure tense marker, synchronically its modal semantics are too strong to ignore.

Apart from this irrealis category, then, two main formal paradigms can be found. One of them, the one exemplified above in (27-9) is prototypically used with present time reference, the other with past time reference. Two differences can be found between them. On the one hand, a different stem of the verb is used: the stem used with past reference generally adds segmental phonological material to the

[9] In the terms of Unruh et al. (2003: 208), the Enenlhet ibi-locative is used to refer to events which originated in a different place from where the speaker is. In this case, the language teacher who this sentence stems from taught me how I would have to say 'I want to learn Sanapaná'. Since I came from a different country to do so, the use of this locative category seems to make sense here. Nevertheless, the exact semantics of Sanapaná locative and directional categories are not the object of this study, and are not relevant to the research objectives of this thesis.

present stem (different increments are attested, and suppletion is also not uncommon), which leads Unruh et al. (2003) to designate it with the term *amplificación radical*, taken over here as *stem extension*. On the other hand, for most verb classes, this extended stem takes a different suffix than the regular stem. Example sentences (35a-c) and (36a-c) show the difference between these two forms of the Sanapaná verb for 'to sleep': the first stem (typically used with present time reference) *jetn* takes a suffix *-eje*, the second stem (typically used with past time reference) has the form *teŋ* (i.e. a case of suppletion), and takes the suffix *-ke*.

(35) Sanapaná (FN2.123)

a. *as-jetn-eje*
1SG-sleep-REAL.V1
'I am lying down/sleeping'

b. *ap-jetn-eje*
2/3M-sleep-REAL.V1
'you (M.SG) are/he is lying down/sleeping'

c. *an-jetn-eje*
2/3F-sleep-REAL.V1
'you (F.SG) are/she is lying down/sleeping'

(36) Sanapaná (FN2.123)

a. *as-teŋ-ke*
1SG-sleep.STX-IPFV.V1
'I lay down/slept'

b. *ap-teŋ-ke*
2/3M-sleep.STX-IPFV.V1
'you (M.SG)/he lay down/slept'

c. *an-teŋ-ke*
2/3F-sleep.STX-IPFV.V1
'you (F.SG)/she lay down/slept'

These forms also confirm the existence of several verb classes with different morphological behaviour in Sanapaná. The suffix *-ke* occurs in combination with the second stem *teŋ* of 'to sleep,' is usually used for past time reference, and contrasts with the suffix *-eje* which is typically used for present time reference. We saw in example sentences (27a-f) that this same *-ke* suffix is also combined with the first stem *taw* of the verb 'to eat,' where it is usually used with present time reference.

Furthermore, the semantics of these suffixes and stems, too, do not only contain an element of time reference. Just as the previous paradigm, the paradigm where the first stem of the verb is combined with a suffix (in this case *-eje*) can, in my view, best be described as a modal category. In (37a-b), for example, it can be seen that the verb forms which were used with present time reference in (35) can also have past time reference when they occur in certain common syntactic combinations with adverbials and/or particles. I therefore follow Unruh et al. (2003) in analysing this first stem and the suffix it combines with as a modal category referring to states of affairs that actually took place in the real world. The temporal component associated with this state of affairs has to be specified through further overt constituents in the syntax, and only if these are absent, the 'default' temporal load of these forms (i.e. 'present') becomes relevant.

(37) Sanapaná (FN2.123)

a. *as-jetn-eje=łke* *sosokhoje*
1SG-sleep-REAL.V1=IMM morning
'I slept this morning'

b. *ap-jetn-eje=łke* *sosokhoje*
2/3M-sleep-REAL.V1=IMM morning
'you (M.SG)/he slept this morning'

The second stem and suffix as well can be argued to carry a greater semantic load than mere tense, even though my evidence for this is more scant. In (38), the form *ejanejaŋka* receives a present habitual reading. The *-aŋka* here is the suffix combining with the second verb stem, and expresses the extensive aspect and the subjunctive (one of the categories of subordination).

(38) Sanapaná (FN2.125)

...tenjo *e-jan-ej-aŋka* *metaʔa* *koʔo*
…then 1SG.O-invite-STX-EXT.IPFV.SBJV always PRO:1SG
'…and you always invite me.'

The habitual aspectual load can be gleamed from the 'extensive' component, which typically refers to complex events that are not delimited in time (Unruh et al. 2010: 176). Nevertheless, the fact remains that the past form (recognisable through the stem extension *-ej*) is used to refer to a habit in the present, implying that the semantic load of this paradigm is not limited to tense. Instead, I rather believe it is one of imperfective aspect: this is both compatible with the present habitual reading which it receives in (38), and it accounts for the fact that this paradigm, when conveying past time reference, rarely occurs with a temporal adverbial. Since it expresses imperfective aspect, it can be argued to be rather vague and unbounded in its default temporal semantics, as opposed to the realis paradigm which, in combination with adverbials or particles with past time reference, refers to a precise, bounded time interval in the past in which a specific event took place.

In sum, I showed in this section that Sanapaná verbal cross-referencing makes a distinction between first person and non-first person arguments, and that within the latter, masculine arguments are treated differently from feminine ones. A hierarchical argument marking system (1 > 2/3) is used in monotransitive clauses. Furthermore, I presented the three main morphologically expressed verbal TAM categories, and argued that the labels 'past,' 'present,' and 'future' tense are too narrow for them since all of them can be used for reference to a moment or period in time different from what is implied by these names. Instead, I proposed treating the categories typically used with present and future time reference as modal categories, expressing actual events (realis) and hypothetical or desired events (irrealis), respectively. The category most often expressing past time reference, then, I attributed mainly imperfective aspectual semantics. It can, however, also be used to express present habitual situations.

Chapter 4 Temporal Predicative Particles in Sanapaná: A Description

With all the necessary theoretical preliminaries being laid out in the two previous chapters, I now move on to the main descriptive part of this thesis. In the three subsequent chapters, I will discuss the formal (both morphophonological and syntactic) and functional properties of three discourse particles in Sanapaná: *ałta* (4.1), *łke* (4.2) and *hata* (4.3).

4.1 ałta

In its prototypical use, the particle *ałta* has a temporal semantic load. In particular, it locates the event expressed by the clause in the prehodiernal past. When used in this way, it occurs in the second position of the clause, immediately following the main verb. *Ałta* can follow both the realis form of the verb (39a), and the imperfective form (characterised by the stem extension as in 39b). As mentioned in chapter 3, the possibility to combine realis verbs with *ałta*, in which case they receive past time reference, was one of the arguments supporting the analysis of these forms as expressing realis mood rather than present tense. In (39a), where this is the case, *ałta* again locates the event expressed by the clause in the past, before the day the utterance is produced.

(39) Sanapaná (FN2.88; 1.125)

a. *tenjo* *as-meja-kh-eje=łta* *nhan* *kajwak* *naʔa* *kánsa*
then 1SG-go-REP-REAL.V1=PHOD CONJ PRO:1SG PRES field.SP
'Then I went to the football field as well.'

b. *as-weta-j-ʔa=łta* *nem=ałta*
1SG-see-STX-IPFV.V1=PHOD sun=PHOD
'I saw you/him/her/it/them yesterday'

As can be seen in (39a-b), *ałta* takes part in a morphophonological process where the first *a* segment is lost when the preceding word ends in a vowel. It is the case in Sanapaná that many common verbal suffixes end in a vowel: *-eje*, *-ke*, and *-ʔa*. The most common exceptions are the irrealis suffixes *-ok* and *-ek*, which, because of their semantics (often containing future temporal reference), are incompatible with *ałta*. Therefore, the most commonly encountered form of this particle is actually *łta,* not *ałta*. This process of elision of an initial syllable (especially a personal prefix) colliding with a word-final vowel is relatively common in Sanapaná (see for example the form *nhan* < *kanhan* 'also' in 39a, and the loss of the initial syllable *ak-* in *nem=ałta* < *aknem=ałta* , 'sun/day-PREHOD' in 39b). Because of this, and

since the full form *ałta* is still clearly visible when it follows a consonant (e.g. in *nem=ałta* in 39b), I will still treat *ałta* as the base form of this particle, and *łta* as a phonologically conditioned allomorph.

Another phonetic particularity of the particle *ałta* is that, in rapid speech, the lateral fricative can be replaced by a glottal stop. As such, *ałta* can in certain contexts (when following a vowel-final word in rapid speech) be reduced to *ʔta* (see, for example, 40a-b). This is, however, a case of free variation and recognised as such by the native speakers.

(40) Sanapaná (FN1.125; 2.19)

a. *as-jetn-eje=łta* *nem=aʔta*
1SG-sleep-REAL=PHOD sun-PHOD
'I was sleeping yesterday'

b. *as-taw-ke=ʔta* *nem=ałta*
1SG-eat-REAL=PHOD sun-PHOD
'I ate yesterday'

This particle can combine with more lexical classes than just verbs, however. In (40), it can already be seen that *ałta* follows the noun *aknem* 'sun/day,' generating the lexicalised meaning 'yesterday'. Apart from this lexicalisation, it is a productive phenomenon as well. In (41a), for example, *ałta* follows the O argument *temakha* (diachronically lexicalised from the copula *ntahak* 'to be', see Unruh et al. 2003: 244, but synchronically functioning as a noun meaning 'manner'). It can even be applied to Spanish loan words, such as (41b) where it follows the adverbial *semana* 'week'.

(41) Sanapaná (FN2. 87; 2.49)

a. *o-łen-ek* *temakha=łta* *nhan* *aknem=ałta* *domíŋgo*
1-speak-IRR manner=PHOD CONJ sun/day=PHOD Sunday.SP
'I am going to speak about yesterday, Sunday.'

b. *eskwéla bíblika de weráno* *pa-j-ʔo* *semana=łta* *lúnes*
school_biblical_summer_school.SP start-STX-SBJV week.SP=PHOD Monday.SP
'The biblical summer school started last week Monday.'

It is clearly visible in (41) that the function of *ałta* is not the same when it follows a noun as when it follows a verb. In this particular example, the main verb *ołenek* is marked for irrealis mood and has future time reference, rather than the past time reference which would be expected if *ałta* had scope over the clause as a whole. Instead, it seems to only have scope over the word immediately preceding it: only the reference of *temakha* is located in the past (a literal translation of the sentence would be something like 'I am going to tell you the past manner of yesterday, Sunday').

When co-occurring with nouns, *ałta* can also have a different function: in (42) below, it does not locate the reference of the noun *ʔawa* 'leaf' in the past: it does not mean 'the things that were leaves,' as one would expect if *ałta* had the semantics of a classical nominal tense. Instead, in this clause (where the speaker talks about how one goes about hunting ostrich), it places the time at which the referent was introduced in the discourse in the past, meaning 'the leaves about which we talked a while ago'.

(42) Sanapaná (FN2.70)

apk-eł-peŋkan-m-ałk-a *ja:met* *ʔawa=łta* *apk-a:tek*
2/3M-PL-put-STX-PAS-IPFV branch leaf=PHOD 2/3M-head
'They put the branches and leaves (which we talked about) on their head.'

Futhermore, *ałta* can co-occur with auxiliaries, and with conjunctions. In (43a), it follows the auxiliary *moʔo*, locating the irrealis modal load (in this case a more specific deontic load of desire) in the past:

"before yesterday, I had a desire to speak." In (43b), it follows the conjunction *tenjo* 'then, subsequently,' and it seemingly has scope over the whole clause, locating it in the prehodiernal past (this sentence came from a story about how the Sanapaná used to fish).

(43) Sanapaná (FN2.140; 2.74)

a. *moʔo=łta* *o-łen-ek* *aknem=ałta*
1SG.AUX.IRR=PHOD 1-speak-FUT sun/day=PHOD
'I wanted to speak yesterday.'

b. *tenjo=łta* *e-takł-e* *ap-ka:tek*
then=PHOD 2/3M-bite-IMP 2/3-head
'Then you had to bite its head.'

(44) Sanapaná (FN1.61)
as-teŋ-ke=łta *en-enta:men=ałta*
1SG-sleep.STX-IPFV.V1=PHOD IMPERS-past=PHOD
'I was sleeping in the past'

Example (44) then, shows the last function of *ałta* that I wish to discuss in this section. The particle here follows the noun (used here as an adverbial) *enentamen* 'past'. It can hardly be said that *ałta* locates the reference of the noun in the past, since the past meaning is inherent in the semantics of this noun itself. Neither can it be said to directly locate the event expressed by the verb *asteŋke* in the past: this is done by the *ałta* directly following this verb. Instead, it seems to emphasise the temporal reference contained within the semantics of the adverb (possibly the reason for Gomes 2013: 113 glossing it as a topicaliser). A more literal translation of (44) would thus be something along the lines of 'it was in the past that I was sleeping.'

It can, however, be said that, rather than a form of emphasis, a more suitable analysis of the function of *ałta* is that it "specifies an existing predicativity" (as Kalisch 2009: 119 argues for Enlhet). In fact, every phrase which co-occurs with *ałta* automatically receives a predicative interpretation. For verbs, this is not surprising: in example (39b), repeated below as (45a), the verb *aswetajʔa* functions as a predicate whether the particle *ałta* follows it or not. Consequently, in this case, *ałta* does not make a predicate of the constituent it is added to, it rather temporally specifies the already existing predicativity of the verb, locating it in the past.

(45) Sanapaná (FN1.125; 2.49)

a. *as-weta-j-ʔa(=łta)* *nem=ałta*
1SG-see-STX-IPFV.V1=PHOD sun-PHOD
'I saw you yesterday'

b. *eskwéla bíblika de weráno* *pa-j-ʔo* *semana=łta* *lúnes*
biblical_summer_school.SP start-STX-SBJV week.SP=PHOD Monday.SP
'The biblical summer school started last week Monday.'

When this particle is added to different word classes, such as the nouns *semana* in (41b), repeated here as (45b), or *enenta:men* in (44), it also triggers a predicative reading. Rather than 'the past' *enenta:men=ałta* means 'to be in the past' resulting in a literal translation of (41) as 'I was sleeping, it was in the past.' Similarly, *semana=łta* does not mean 'last week,' but rather 'it was last week,' yielding a literal translation of (45b) as 'The biblical summer school started, it was last week Monday.'

It was established in section 3.3 above that nouns in Sanapaná can optionally receive a predicative reading even when they are not used in combination with a particle such as *ałta*: *semhen* can either mean 'a dog' or 'to be a dog,' *as-tawa* can mean 'my spouse' or 'she is my spouse.' They thus possess what Kalisch (2009) calls inherent predicativity. As was the case for verbs, the syntactic combination of a

noun with *ałta* does not make this noun predicative, but it specifies its predicativity, or makes it explicit: when combined with this particle, these nouns lose their referential reading and always function as a predicate.

4.2 łke

The second particle under discussion in this thesis, *łke* (allomorphs *łkeʔek, ʔkeʔek, łkeʔe* and *ʔke*), has more or less the same syntactic characteristics as *ałta*. That is to say, it can combine with verbs, nouns, auxiliaries, and conjunctions, and always follows the word it modifies. example (46a) shows the prototypical use of this particle in combination with a verb:

(46) Sanapaná (FN2.89)

a. *tenjo sosokhoje-na naʔak e-jawek-aŋk-okh-a=łke Jens*
then morning-LOC PRES 1SG.O-call.STX-CPL-ABIT[10]-IPFV.V2=IMM Jens
'Then this morning, Jens called me'

b. **tenjo aknem=ałta e-jawek-aŋk-okh-a=łke Jens*
then sun/day=PHOD 1SG.O-call.STX-CPL-ABIT-IPFV.V2=IMM Jens
'Then yesterday, Jens called me'

As can be seen in the translation of (46a), when *łke* follows a main verb, it also contributes past temporal semantics to the clause. *łke* diverges from *ałta*, however, in that it refers to a hodiernal past and as such complements the latter. This can be seen in (46b): here, the adverbial *sosokhojena* 'this morning' is replaced by *aknem=ałta* 'yesterday,' which makes the use of *łke* in combination with the main verb ungrammatical. In this context, only the use of *ałta* instead of *łke* results in a grammatical sentence.

Just as demonstrated in 4.1 for *ałta*, *łke* can combine both with the realis stem of verbs (as is the case in 47a below), and with their imperfective stem (as in 46a and 47b). In both cases, it locates the event expressed by the verb in the hodiernal past, providing additional evidence for the analysis of these categories as modal and aspectual rather than temporal, since they can both easily be used to achieve the same temporal reference.

(47) Sanapaná (FN2.123; 2.140)

a. *as-jetn-eje=łke*
1SG-sleep-REAL.V1=IMM
'I was sleeping (earlier today).'

b. *as-teŋ-ke=łke*
1SG-sleep.STX-IPFV.V1=IMM
'I was sleeping (earlier today).'

[10] Unruh et al. (2003) label the *-okh* suffix as *abitive*. It is used both as a directional suffix, indicating that the event is proceeding in a direction away from the speaker; and as a category of quantitative aspect, indicating that the event expressed by the verb is complex (i.e. consists of various subevents), and an instance of an event that takes place habitually. In this case, I habitually called this language teacher during my stay in La Esperanza, and at this specific instance I had to try multiple times before he took up his phone. The semantics of this category thus seem to check out at first sight in Sanapaná. More research is necessary, however, to disentangle the forms and functions present in complex verb forms such as *ejawekaŋkokha* here.

When combining with an auxiliary, *łke* functions rather differently than *ałta*, however. It still follows the irrealis auxiliary *maʔa* as in (48a), but the semantic load it contributes is not the same. It does not locate the irrealis aspect expressed by the auxiliary in the past (which would result in a meaning such as 'I was going to drink'). Instead, it seems to contribute in these cases a prospective aspectual load ('I am going to drink right away').

(48) Sanapaná (FN1.20; 1.21)

a. *maʔa=łke* *o-jn-ek*[11]
AUX.IRR=IMM 1-drink-IRR
'I am going to drink. (For example, responding to 'Do you want a drink?')'

b. *maʔa=łke* *o-tjep-ok* *sokhoje* *sosokha*
AUX.IRR=HOD 1-leave-IRR moment tomorrow
'I need to leave for a brief moment tomorrow.'

Example (48b) shows that *łke* does not necessarily have to temporally locate an event on the day of the utterance when it is combined with an irrealis verb (as opposed to when it is used with past time reference). Here, the adverbial *sosokha* 'tomorrow' locates the event in the post-hodiernal future (on the day after the utterance). Nevertheless, it can be said that the particle in this case still expresses a prospective aspect. The comment in (48b) was made in a context where I had just asked this consultant whether he would be willing to work more with me over the course of the next two months. He replied that he did, but he would not be able to work the next day because he had to leave the community for a while. Therefore, against the backdrop of the two-month period that was the focus of attention, the *leaving* event taking place *tomorrow* can still be seen as an immediate future.

Since it is clear that *łke* does not convey pure temporal reference (it can refer to events both in the past and in the future), I here analyse it as an aspectual marker expressing both prospective and retrospective aspect (i.e. taking a point of view immediately prior or posterior to the event expressed in the clause). This is the reason why I gloss it as IMM for *immediate aspect*, subsuming the two aforementioned aspectual categories (in the sense of Givón 2001: 298-300).

Additionally, *łke* can syntactically combine with nouns. When doing so, its function is parallel to that of *ałta*: it makes reference to the moment when the noun it follows was introduced in the discursive context. The two particles enter in opposition with each other in that *łke* locates the introduction of the referent in the recent discursive context (as shown in 49a), whereas *ałta* is used for referents that were introduced longer ago. Additionally, just as *ałta*, *łke* triggers a predicative reading of the noun it combines with. The noun *nemakha*, for example, can mean both 'house' and 'be a house' when used in isolation, but the use of *łke* triggers the latter reading, resulting in a literal translation of (49a) as 'I am going to return (to that place). It is my house.'

(49) Sanapaná (FN1.65)

a. *moʔo* *o-ta-kh-oho* *as-nemakha=łkeʔe*
1.AUX.IRR 1-go-ABIT-INTNS 1SG-house=IMM
'I am going to return to my house (that I just mentioned) again'[12]

[11] The reason why in these examples the auxiliary takes the form *maʔa* rather than *moʔo* is not as yet completely clear to me. As of now, it seems that when the auxiliary precedes the main verb immediately, it must agree with the A argument of the clause (resulting in the form *moʔo* for first person singular), whereas agreement is optional when auxiliary and main verb are separated.

[12] In all the English free translations of examples (49, 50, 57, 59, 60, 61, 65, 70, 71, 76, 77) 'just' is used in the sense of 'recently, a moment ago', rather than 'only, merely'.

b. *kełwoje=łke* *moʔo* *o-w-ata*
moment=IMM 1.AUX.IRR 1SG-arrive-VENT.IRR
'I will soon arrive (back) here.'

This becomes clearer in (49b). Here, *łke* combines with the intensive form (recognisable through the suffix *-oje*) of the noun *kełwa*, meaning 'moment.' This form *kełwoje* is a commonly used adverbial meaning 'soon, right away.' As such, it already conveys prospective aspectual semantics without the presence of *łke*. The function of the particle in this case, then, is purely to provide the noun with a predicative, topicalised reading, resulting in a direct translation such as 'It will be soon, I will arrive back here.'

Ultimately, I wish to make some observations about the form of this particle. It can be argued that its full, original form is *łkeʔek* rather than *łke*. This form can be seen in example (50a) from Unruh & Kalisch (2003). Nevertheless, in my own corpus I only encountered one instance of this form (50b), and even in this example the lateral fricative has been reduced to a glottal stop. The closest formal manifestiation of this particle found with any frequency in my corpus is *łkeʔe* (as in 49a above), with loss of the final *-k* (which, as noted before, is a common process in Sanapaná).

(50) Sanapaná (Unruh & Kalisch 2003: 4; FN2.68)
a. *as-m-eje=łkeʔek*
1SG-buy-REAL.V1=IMM
'I just bought it'[13]

b. *ask-elh-tenna-s-ke=ʔkeʔek*
1SG-PL-tell-STX-IPF.V1=IMM
'I was just been telling (stories…)'[14]

As can be seen throughout this section, the most commonly found form of the particle is simply *łke*, with complete loss of the last syllable. This could be the result of the advancing phonetic erosion of intervocalic glottal stops in Sanapaná (also present, for example, in the future auxiliary *moʔo* at times sounding more like *mo:*), and the shortening of the resulting final long vowel. Even though there seems to be a preference for users to pronounce the longer form *łkeʔe* before a pause and *łke* clause-internally (see 49a for the former and 48a-b for the latter), this is not a hard-and-fast rule. The shortened form *łke* also frequently occurs in clause-final position (as in 47a-b). I do not have any examples, however, of the long form *łkeʔe* being used clause-internally.

(51) Sanapaná (FN1.18)
as-taw-ke=ʔke *sosokhoje*
1SG-eat-REAL.V1=IMM morning
'I ate this morning'

Lastly, in rapid speech a phonological process similar to that affecting *ałta* can further erode the segmental form of this particle: the replacement of the lateral fricative by a glottal stop. This is seen in examples such as (50b) and (51) above. As a consequence of the two processes laid out here, the particle under discussion in this section can occur as a wide range of forms depending on its position in the

[13] I replaced Unruh & Kalisch' (2003) segmentation *as-m-e ye* (1SG-buy-REAL REAL), with *ye* as a 'factive' particle, by *as-m-eje* (1SG-buy-REAL). Unruh & Kalisch (2003) see *ye* as an obligatory particle expressing the 'factive' (my 'realis') category. In my examples throughout this thesis, however, it is shown that not all realis forms take this particle. Therefore, I analyse it as part of the realis suffix of some verb classes.

[14] The reason for the verbal plural marker *eł-* on this verb is that, apart from indicating the number of A or S, it can also express pluralisation of O or of the event as a whole (which is the case here, see footnote 4 p.35).

clause, the speed of speaking, and individual speaker preference: *łkeʔek*, *ʔkeʔek*, *łkeʔe*, *łke* and *ʔke* are all attested. Whereas for *ałta*, there were clear contexts where the full form always surfaced, this is not the case for this particle. Therefore, I will in this thesis use the most common form *łke* when referring to the complete set of five forms this particle can take on.

4.3 hata

The last particle to be discussed in this thesis is *hata*, which can in several respects be seen as the mirror image of *ałta*: it has the same syntax and generally the same functions, except for expressing future rather than past time reference. In its most prototypical use, it again follows a clause-initial main verb, as is the case in (52a-c). This main verb is most often marked for irrealis. As we saw in 3.4.2, irrealis-marked verbs typically receive a deontic modal reading (for example, 'I want to sleep more'). The addition of the particle *hata* seems to change this to a more epistemic reading (specifically, a prediction, which lies semantically closer to a future tense). It can even present the proposition as non-modal altogether, conveying a purely temporal future reading.

(52) Sanapaná (FN2.83, 2.123, 2.138)

a. *o-l-weta-ł-ek=hata* *sososkha*
1-PL-see-REFL(?)-IRR=FUT tomorrow
'We will see each other tomorrow'[15]

b. *o-tjen-ek=hata* *mokham*
1-sleep-IRR=FUT more/again
'I am going to sleep more'

c. *heŋkaʔe* *talhnaʔmoho* *naʔak* *o-ta-ʔakha=ta*
DEM.PROX afternoon PRES 1SG-go-DIR=FUT
o-nj-akha=ta *as-kawa-kha*
1SG-go-DIR=FUT 1SG-country-LOC
'This afternoon I will leave and go to my country'

Hata can make reference to a posthodiernal future (as in 52a), or an unspecified future (as in 52b) in which it shows similar behaviour to *ałta*. Nevertheless, *hata* can also be used in clauses which are explicitly located in the hodiernal future, something *ałta* cannot do in the past (52c).

Interestingly, when *hata* co-occurs with an irrealis-marked verb which is preceded by the irrealis auxiliary (53), the particle follows the main verb and does not interrupt the Aux-V sequence as *łke* and *ałta* do. Here as well, *hata* seems to diminish the modal semantics of volition usually conveyed by the auxiliary in favour of a purely temporal reading, presenting the event as more certain than when *hata* is not used.

[15] The stem of the verb 'to see' is *weta*, and the irrealis suffix is *-ek*. I do not know what the status is of the *ł* in *olvetałek*. It could have a linking function between the two vowels, although in that case it would be expected for the *-e* to drop and result in *olvetak*. Therefore, I hypothesise that it is a reciprocal marker, since the reciprocal semantics of this example otherwise go unaccounted for. I have not found such a reciprocal marker anywhere else, however, making this hypothesis very preliminary.

(53) Sanapaná (FN2.122)
maʔe *e-tjen-ek=hata*
2/3M.AUX.IRR 2/3M-sleep-IRR=FUT
'You/he is going to sleep'

In combination with nouns and conjunctions, *hata* functions in the same way as *ałta*: it prompts a predicative interpretation of this constituent. In (54a), for example, *hata* co-occurs with the conjunction *tenjo* 'then'. What *hata* does, here, is that it triggers a reading of *tenjo* as a predicate 'to be later'. In addition, even though the future time reference contributed by *hata* is in part obscured by the fact that *tenjo* itself already inherently refers to the future, the particle still locates this predicate in the future (we saw in 46a that *tenjo* can also introduce clauses with past temporal reference). The literal translation of (54a) would then be 'It will be later. We will eat watermelon.'

(54) Sanapaná (FN2.85, 1.47)

a. *tenjo=ta* *o-ł-t-oʔok* *sammammehe*
then=FUT 1-PL-eat-IRR.PL watermelon
'Then, we will eat watermelon'

b. *pesa:sep=hata* *neła* *nhan* *en-maʔtma-ma* *polisia* *nenłet* *anaʔak*
night =FUT maybe CONJ PL-work-INF police.SP Sanapaná PRES
'At night maybe the Sanapaná police will be working'

In (54b), *hata* combines with the noun *pesa:sep* 'night'. Here, both of its functions are clearly visible: on the one hand, it triggers a predicative reading of the preceding noun, i.e. meaning 'to be night'. Then, it locates this predicate in the future (the speaker is talking here about what will happen on New Year's Eve, which was at the point of this recording more than a week away). All these things taken together, the literal meaning of (54b) can be represented as 'It will be night. Maybe the Sanapaná police will also be working.'

Lastly, I briefly discuss the segmental form of this particle, as I did for the two previous ones. The attentive reader will have noticed already that the particle glossed as FUT comes in two different forms in the example sentences above. Once again, which form is used depends on the final segment of the preceding word, and on the speed of speech. Just as the first *a* of *ałta* is often lost when it follows a vowel-final word, so is the first syllable *ha* of *hata*. This is seen in examples (52c) and (54a), where *hata* follows respectively the final *a* of a verb, and the *o* of *tenjo*.

Furthermore, in rapid speech, *ha* can also be lost when *hata* follows a *k*, in which case this *k* itself is lost as well. As mentioned before, simultaneous elision of a clashing word-final *k* and a word-initial vowel is relatively common in Sanapaná, but in this situation it also affects the word-initial *h*. This is often the case when *hata* follows an irrealis-marked verb, such as in (53), which in connected speech easily surfaces as (55).

(55) Sanapaná (FN2.122)
maʔe *tjen-e=ta*
2/3M.AUX.IRR sleep-IRR=FUT
'You are/he is going to sleep'

Now that both the formal and functional aspects of the three particles under scrutiny in this study have been described, the next two chapters will provide a more grounded discussion of these aspects, both with regards to the theoretical framework of the analysis of TAM categories laid out in chapter two, and with regards to the other languages of the Enlhet-Enenlhet family.

Chapter 5 Temporal Predicative Particles in Sanapaná: A Theoretical Discussion

In this chapter, the descriptive findings concerning the particles *ałta*, *łke* and *hata* will be discussed from a theoretical and comparative perspective. This theoretical discussion of Sanapaná *ałta*, *łke* and *hata* will consist of three separate sections. Firstly (5.1), I will briefly present some evidence for their formal status as particles, rather than bound morphemes (i.e. inflectional or derivational affixes) or adverbs. Subsequently, I move on to apply the theoretical concepts introduced in chapter 2 to the Sanapaná data. I first explain the temporal reference these particles convey in formal terms (5.2); and then discuss their implications for the analysis of Sanapaná as a language with flexible word classes (5.3).

5.1 The status of ałta, łke and hata

In this first section, I would like to briefly justify my decision to treat *ałta*, *łke* and *hata* as separate particles rather than bound morphemes or adverbs. Let me first argue why these morphemes, in my view, should not be treated as adverbs, even though some of their meanings are compatible with that which one would expect from adverbs of time (e.g. *łke*, which can in contexts such as (48a), repeated here as (56), be translated as 'right away').

(56) Sanapaná (FN1.20)
maʔa=łke *o-jn-ek*
AUX.IRR=IMM 1-drink-IRR
'I am going to drink (right away).'

Firstly, it is not clear that Sanapaná even has a class of adverbs at its disposition. In the semantic domain of time, in any case, adverbial constituents typically simply consist of a noun: *pesa:sep* 'night,' for instance, can function as an argument of a verb ('a night'), as a predicate (meaning 'to be night'), or as a time-indicating adverbial (meaning 'at night, last night'). Additionally, adverbs are usually seen as lexical items: adverbs of time would then express a lexical meaning, such as English *later*, *yesterday*. The Sanapaná particles are distinct from such canonical adverbs in that they have, next to the lexical temporal meaning they possess, also a purely grammatical meaning: they make explicit the inherent predicativity of the word that they follow.

Furthermore, if a language has a separate lexical class of adverbs, they can usually function as independent phonetic words, with their own stress pattern. The Sanapaná particles, on the other hand, can never be independent phonetic words. Instead, they always encliticise to the preceding word, forming

together one phonetic unit. In this process, stress shifts away from the root and onto the clitic: as such, forms such as *enenta:men'ałta* (from *enen'ta:men* 'past'), *kełwoje'łke* (from *kełwo'je* 'moment') and *pesa:sepha'ta* (from *pe'sa:sep*) are created.

If they are not adverbs, might they then be bound morphemes? The temporal meaning which they express makes them, at first sight, plausible candidates for an analysis as tense or aspect suffixes. Nevertheless, I believe an analysis of these morphemes as inflectional suffixes does not hold either, since *ałta*, *łke* and *hata* are all optional: all contexts where they appear would be equally grammatical without them. This contrasts with other candidates for the position of inflectional affixes, such as the person markers (e.g. *as-* and *ap-*) and the realis/irrealis mood markers (such as *-ek* and *-eje*) which are obligatory (except for when parts of them are elided due to predictable morphophonological processes). Furthermore, *ałta*, *łke* and *hata* are applicable to members of almost all word classes, including conjunctions (such as *tenjo* 'then') which do not take any other inflectional morphology.

That these clitics are not derivational morphemes either, can be intuitively hypothesised from the cross-linguistic generalisation that tense and aspect are (virtually) always expressed either inflectionally or analytically (i.e. through auxiliaries and other verb-external material). Nevertheless, Bybee and Dahl (1989: 55) make the argument that (at least in some languages), the expression of iterative and inceptive aspectual categories should be classified as derivation rather than inflection. Therefore, further argumentation is needed to exclude this possibility for the Sanapaná particles.

With this respect, the last argument used against their analysis as inflectional categories – their applicability to members of all word classes – also impedes an analysis of these morphemes as derivational affixes. According to Booij (2006: 655), derivational affixes typically apply to members of one specific lexical class in order to convert them into members of a different class, such as the suffix *-er* in English, which can be applied to verbs (e.g. *walk*) to convert them into nouns (*walker*), but cannot be applied to, for instance, nouns or adjectives. The morphemes *ałta*, *łke* and *hata* can all be applied to multiple word classes, and do not form new lexemes by changing the word class of the root they are applied to. Furthermore, Booij (2006: 659) argues that derivational morphemes occur closer to the word stem than inflectional ones, since derivational morphemes create a new lexeme which then needs to be inflected before its insertion into syntax. In Sanapaná, however, these three morphemes always occur on the boundary of a word, after all inflectional morphemes. For these reasons, I argue that an analysis of *ałta*, *łke* and *hata* as derivational morphemes is also untenable.

In sum, this section laid out phonetic, semantic and syntactic arguments against the analysis of *ałta*, *łke* and *hata* as adverbs, inflectional suffixes, and derivational suffixes. An analysis as adverbs is undesirable since they are clitics rather than independent words, and since they contain purely grammatical in addition to lexical information. They can hardly be seen as inflectional morphemes, since they are optional; or derivational suffixes, since they are more peripheral than inflectional affixes and do not change the word class of their host. As a consequence, I propose to treat the three morphemes concerned here as clitics which are syntactically independent, but phonetically depend on the immediately preceding word, warranting the use of the term *particle* throughout this thesis.

5.2 The temporal semantics of ałta, łke and hata

As was roughly described in chapter 4, the three particles that are the subject of this study all possess a certain semantic load related to temporal reference. In this section, I will discuss this semantic load, and its interplay with verbal TAM categories in more formal terms, inspired by the frameworks on tense, aspect, and modality presented in chapter 2. Subsection 5.2.1 will do this for the particles in combination with verbs, subsection 5.2.2 treats the situations when the particles follow nouns.

5.2.1 Verb-particle combinations

The first particle, *ałta*, when used in combination with a verb, specifies the temporal location of the event expressed by this verb. To be precise, it locates this event in the period before the day on which the utterance was produced. In Reichenbach's terms, one could thus say that it posits the day of the utterance as the reference time R with respect to which the event is located, and that it locates the event time E expressed by the verb before this: E > R(S).[16] Alternatively, one could say that *ałta* expresses a situation where R and S coincide, and both are preceded by E; and that it additionally invokes a remoteness condition stating that E must occur at least one day before R and S.

Secondly, *łke* posits a bit more of a challenge. It portrays the event expressed by the verb it follows either as just finished (on the day of the utterance), or about to begin (on the day of the utterance or later, when a larger timeframe is being used). As such, it can be said that *łke* either locates the final boundary of E right before R (without taking into account the beginning of E), or it locates the initial boundary of E right after R (disregarding the end point of E). There are thus two possible timelines for the temporal reference of *łke*: the former scenario can be represented as E2 > R; the latter one as R > E1.

Hata, ultimately, is similar to *ałta* and differs from *łke* in that it expresses the temporal location of the event expressed by the verb it follows rather than focusing on one of the boundaries of this event. In contrast with *ałta*, however, it locates this event in the future, regardless of whether it occurs later on the day of speaking or after this day. Therefore, the relative time line of E, S, and R looks as follows for *hata*: S, R > E. Here, R coincides with S instead of including it. The event is located relative to the (instantaneous) moment of the speech act, rather than to an interval starting at the beginning of the day of the speech act, and finishing at the end of this day. Another way to put this is that *hata*, as opposed to *ałta*, does not encode a remoteness distinction. In this respect, Sanapaná is in agreement with De Haan's (2011: 450-1) hypothesis that languages typically make more remoteness distinctions in the past than in the future.

According to the definitions assumed in chapter 2, one must thus necessarily interpret *ałta* and *hata* as expressing tense, since they express the temporal *location* of the whole event E with respect to a (coinciding) R and S. The former then expresses a prehodiernal past tense, the latter a future tense without remoteness component. On the other hand, *łke*, must be seen as an aspect marker, since it represents the temporal structure of an event rather than its location. In particular, it is a marker of phasal aspect, since it emphasises one specific stage of an event. Since these stages in this case are the moments right before the beginning or right after the ending of E, it can be said that *łke* combines the meaning of a

[16] The brackets in this formula mean that S is included in R, i.e. R is a time interval (consisting in this case of a whole day), and S lies completely within the boundaries of R.

prospective aspect marker and a resultative aspect marker, meanings which are subsumed under the term *immediate* aspect (Givón 2001).

The semantic loads of these particles can thus be said to complement the TAM categories expressed on the Sanapaná verb rather well. As argued in chapter 3, many verbal TAM categories in fact show a mix of temporal and modal semantics, often being interpreted as a modal realis-irrealis distinction when they occur in the absence of one of these particles. When either *ałta* or *hata* is added to such a verb, it can be said to resolve this ambiguity in favour of a temporal reading of the verb.

5.2.2 Noun-particle combinations

As shown in the previous chapter, all three particles can also modify nouns. When *łke* is used in this way, I only have examples of it fulfilling a discourse-related function, i.e. expressing that the referent of the noun it follows was introduced in the discourse shortly before. This was the case in (49a), repeated below as (57). As such, it does not fulfil a temporal reference function, since it does not locate the existence or the possession of the referent in time.

(57) Sanapaná (FN1.65)
moʔo *o-ta-kh-oho* *as-nemakha=łkeʔe*
1.AUX.IRR 1-go-ABIT-INTNS 1SG-house=IMM
'I am going to return to my house (that I just mentioned) again'

The other two particles, *ałta* and *hata*, on the other hand, can fulfil such a temporal reference function. This is the case in (41a-b), and (54b), also repeated below for the reader's convenience.

(58) Sanapaná (FN2.87, 2.49, 1.47)
a) *o-łen-ek* *temakha=łta* *nhan* *aknem=ałta* *domíngo*
1-tell-IRR manner=PHOD CONJ sun/day=PHOD Sunday.SP
'I am going to tell about (how it was) yesterday, Sunday.'
b) *eskwéla bíblika de weráno* *pa-j-ʔo* *semana=łta* *lúnes*
biblical_summer_school.SP start-STX-SBJV week.SP=PHOD Monday.SP
'The biblical summer school started last week Monday.'
c) *pesa:sep=hata* *neła* *nhan* *en-maʔtma-ma* *polisia* *nenłet* *anaʔak*
night=FUT maybe CONJ PL-work-INF police Sanapaná PRES
'At night maybe the Sanapaná police will be working'

In (58a), *ałta* explicitly locates the two nouns which it follows in the prehodiernal past: the *temakha* 'manner' which is spoken about is 'the manner in which things happened yesterday,' and the *aknem* 'day' which is spoken about is 'the day before today'. The same goes for (58b), where *semana* 'week' is located in the prehodiernal past by *ałta*, and in (58c), where the *pesa:sep* 'night' that is modified by *hata* is 'a night after today'. In the terms of Tonhauser (2007; 2008), it can be said that *ałta* establishes a relation between a noun phrase time t_{NP} which consists of the entire day of the utterance and a nominal time t_{NOM} which precedes this: $t_{NOM} > t_{NP}$ (S). *Hata*, on the other hand, has a specific moment as its noun phrase time (the moment of speaking), and locates the nominal time after this: $t_{NP,}S > t_{NOM}$.

Is it, therefore, warranted to describe *hata* and *ałta* as nominal tense markers? According to the criteria laid out in Nordlinger & Sadler (2004a: 778), it is not. The two markers under discussion here comply with three of their four criteria (they express a temporal relation, are productive across the whole word class of nouns, and are not restricted to verbless clauses). They fail their fourth and last criterion, however, saying that the possible tense marker needs to be part of the morphology of the noun rather than a syntactically independent element. As demonstrated in the previous subsection, although they

phonetically depend on their nominal host, they do not do so syntactically. As such, they fail this criterion.

A different set of criteria leads us to the same conclusion: *hata* and *ałta* comply with three of Tonhauser's (2008: 333) criteria, but fail to meet the fourth one. Neither of the particles seem to encode a state change (for example, the *hata* in 58c does not necessarily imply that the police will not be working on the day of the utterance), they cannot co-occur (or at least no instances of this are found in my data), and their meaning can be constrained or specified by additional constituents, as in the case of (58b) where *lunes* constrains the meaning of *semana=łta*. Nevertheless, neither particle seems to comply with Tonhauser's fourth criterion, that of anaphoricity. Although it is possible that this is a gap in the data, at the present moment the evidence is in any case too weak to classify *hata* and *ałta* as nominal tense markers.

In sum, *ałta,* and *hata* are classified as verbal tense markers in this thesis, even though they occur outside of the inflectional morphology of the verb. The former expresses a temporal relation that can be summarised as E > R(S); the relation expressed by the latter can be represented as S, R > E. On the other hand, *łke* is analysed as a verbal aspect marker, since it is more concerned with the temporal structure of the event (emphasising the run-up to the event or the immediate aftermath) rather than its temporal location. This part of the semantics of *łke* can be represented as E2 > R / R > E1. *Hata* and *ałta,* as opposed to *łke,* can also convey nominal temporal relationships, respectively $t_{NP,}S > t_{NOM}$ and $t_{NOM} > t_{NP}$ (S). Nevertheless, neither under Nordlinger and Sadler's (2004) nor under Tonhauser's (2008) definition do these particles qualify as nominal tense markers.

5.3 Ałta, łke and hata and the question of flexible word classes

As was demonstrated in chapter 3, nouns and verbs in Sanapaná show a remarkable number of similarities. At the morphological level, both word classes take the same sets of person-marking prefixes, used on verbs to indicate the person and number of the S, A, or O argument, and used on nouns to mark the person and number of the possessor or referent. At the syntactic level, members of both word classes can be followed by the particles *hata, łke* and *ałta*. At the functional level, ultimately, both nouns and verbs can function as predicates, as was also demonstrated in chapters 3 and 4. Specifically, bare nouns can either receive an interpretation as predicates or as referents, whereas nouns followed by one of the three particles automatically receive a predicative interpretation.

These facts prompt the question of whether or not Sanapaná can be seen as a monocategorial language (in the terminology of Evans & Osada 2005), or a language with flexible lexemes (in the terminology of amongst others Hengeveld et al. 2003; Rijkhoff & Van Lier 2013). First and foremost, both frameworks only treat the functional characteristics of word classes. The fact that Sanapaná nouns and verbs are clearly distinct categories based on their morphology (verbs taking aspectual, modal, and directional morphology which nouns do not) is therefore irrelevant to this discussion.

Evans & Osada (2005), as explained in chapter 2, propose three characteristics which a language must adhere to in order to qualify as a monocategorial language, i.e. a language where nouns and verbs together form one syntactic class: compositional semantics, exhaustivity, and (bidirectionally) equivalent combinatorics. Sanapaná seems to meet two of these, but fails the third. The principle of compositional semantics seems to be met. The semantics of a predicatively used noun are straightforwardly

derived from those of the noun itself and those of the predicative function: a noun X used predicatively in Sanapaná receives the meaning ‘to be an X’. This seems to be valid throughout the lexicon rather than only for subsets of it, which makes Sanapaná meet the second criterion as well.

Nevertheless, the third criterion, that of bidirectional distributional equivalence, according to which lexemes must occur in predicative and referential function with roughly the same frequency, is only partly met. Nouns can take on prototypical verbal functions (i.e. predication) without needing any additional morphological or syntactic material: as was argued before, predicative particles such as *hata, łke* and *ałta* only make the predicativity of a noun explicit, they do not make a noun predicative. Nevertheless, the reverse does not hold. In order to be able to function as a referring expression, a verb does need additional morphology: most often the nominalising morpheme (glossed as infinitive) *-ma*. In such cases, the resulting deverbal noun can be interpreted as an object (the one who is X-ed) or as an action (the act of X-ing), which also constitutes a violation of the compositional semantics principle. For example, the deverbal noun *en-to-ma* (1PL-eat-INF) can mean ‘our food’ or ‘our act of eating’.

In the framework advocated by Hengeveld et al. (2003) and Rijkhoff & Van Lier (2013), a language with one flexible class of lexemes which can function both as the head of a referential phrase and the head of a predicate (for which Sanapaná seems to be a candidate) would be classified as a type 1 language in terms of its inventory of word classes. Although the criteria proposed by these authors are less strict, this does not help the Sanapaná case. On the one hand, these authors adopt a less severe form of the semantic compositionality criterion, but since Sanapaná already passed the stricter form of this, this does not seem relevant. On the other hand, they advocate a less strict form of the distributional equivalence criterion: instead of being as frequent in predicative and referential function, it suffices for all lexemes of the major class to have the ability to simply be used both as a referential expression and as a predicate without morphological or syntactic modification, regardless of the frequency with which this occurs Nevertheless, the fact that Sanapaná verbs cannot function as arguments without being explicitly nominalised disqualifies the language from an analysis as a type 1 language as well.

Instead, it seems that Sanapaná must be classified as a language with unidirectional word class flexibility. In such languages, in Beck’s (2013) terms, the functional distinction between nouns and verbs is only neutralised in certain contexts. In the case of Sanapaná, this would be the predicative context: here, nouns and verbs can both occur naturally. In other contexts, on the other hand (such as the argument function), two distinct word classes can still clearly be perceived, since one of them (the verbs) needs an extra morphological operation to be able to occur here. Sanapaná can thus be called omnipredicative (since all lexemes can function as predicates) but not monocategorial (since there are still functional differences between nouns and verbs).

Having discussed the temporal semantics and the role in the word class system of the Sanapaná particles *hata, łke* and *ałta* from a theoretical point of view, it is time to broaden the scope of this investigation. The next chapter will discuss the form and function of the cognates of these three particles in the other languages of the Enlhet-Enenlhet language family, inasmuch as usable data are available.

Chapter 6 Comparison across the Enlhet-Enenlhet Language Family

From a brief glance at the data available on the other languages of the Enlhet-Enenlhet family, it becomes clear that Sanapaná is not the only one which makes use of the set of temporal predicative particles described above. In the first section of this chapter, I will bring data from Angaité (Villagra Carrón p.c. 2017), Enxet (Elliott p.c. 2017), Enlhet (Kalisch 2009), Guaná (Unruh & Kalisch 1999) and Enenlhet (Unruh et al. 2003) on the use of these particles in the picture. Section 6.2, then, is where I take a historical glottometric approach to these data, establishing the subgroupiness of the different subgroups of Enlhet-Enenlhet languages in order to gain insights into the structure of the Proto-Enlhet-Enenlhet dialect continuum. Of course, these insights are only based on the data presented on the three particles discussed in this thesis, and must be supplemented by additional lexical and grammatical data in order to establish these genetic relations more soundly.

6.1 Data

Before applying the methodology of historical glottometry to the Enlhet-Enenlhet languages, I first present data on the form and use of the cognates of Sanapaná *ałta*, *łke*, and *hata*.

6.1.1 Angaité and Enxet

Angaité is the language on which, at the moment, published linguistic data are most scarce. Only two short phrases were found with any relevance for this study. The following two Angaité phrases demonstrate that the language has a cognate of Sanapaná *łke* at its disposition (59a-b).

(59) Angaité (Enlhet-Enenlhet, Unruh & Kalisch 2003: 210-211)[17]

a. *as-m-eʔe* *łkek*
1SG-buy/grab.3.PS-FACT REC
'I just bought/grabbed it.'

[17] The morpheme *-m-* in (56a) is glossed *buy/grab:3.ps* in the source. It is not explained what the gloss *ps* means, but from all the examples in which it is used, I conclude it refers to what Unruh & Kalisch call the *present* stem of the verb (called *realis* throughout this work), which can also, as in the case of (56a), be used with past time reference. The *3* seems to refer to a third person object. The *factive* is the Angaité equivalent of what I call *realis* throughout this thesis.

b. *ap-s-a:kh-e* *łkek*
2/3.M-take_away.3.PS-REP-FACT REC
'He took it away recently.'

These examples unfortunately do not tell us anything more than that a cognate of Sanapaná *łke* exists in Angaité with roughly the same form and meaning. Since it occurs following a vowel and sentence-finally twice, nothing can be said about possible morphophonological processes in which the particle participates; nor does the gloss 'recent past' shed light on its exact semantics.

Furthermore, the examples in (60a-c), courtesy of Rodrigo Villagra Carrón (p.c. 2017), show that Angaité also has cognates of the other two particles.[18] The hodiernal past particle can apparently take the form *łajk* as well, with an open, central vowel rather than the closed front vowel in Sanapaná. The forms of *ałta* and *hata* are, at first sight, the same as in Sanapaná. In addition, (60b) exemplifies the predicativity of word classes other than verbs: *awanhe* is followed by the particle *ałta* and consequently is automatically interpreted predicatively ('the river was big') rather than attributively ('the big river').

(60) Angaité (Enlhet-Enenlhet, Villagra Carrón p.c. 2017)
a. *ap-wu-akt-ek* *łajk*
2/3M-go-VENT-IRR HOD
'you (M.SG)/he just arrived.'

b. *awanhe* *ałta* *alwata*
big PHOD riacho
'The river was big.'

c. *o-ta=hata* *ko:*
1-bring.IRR=FUT PRO:1SG
'I will bring it.'

For Enxet, two phrases were found with the same glosses and free translations as (59a-b), also demonstrating the existence of a cognate to Sanapaná *łke* (61a-b).

(61) Enxet (Enlhet-Enenlhet, Unruh & Kalisch 2003: 210-211)
a. *ek-m-e:k* *łek*
1SG-buy/grab.3.PS-FACT REC
'I just bought/grabbed it.'

b. *ap-s-a:kł-e:k* *łek*
2/3M-take_away.3.PS-REP-FACT REC
'He just took it away.'

In addition, John Elliot (p.c. 2017), who is working on an ELDP-funded Enxet documentation project, generously provided me with some of his insights into the functioning of these morphemes in Enxet. Although he had no examples at hand, he believed it was possible in Enxet to use *łejk*, written <xeyk> in the Enxet orthography developed by the Anglican Church of Paraguay in the 1990s (Kidd 2000: ix), with a prospective meaning, as it is in Sanapaná.

The other two particles, with the forms *ałta* and *sat(e)*, in Elliott's (p.c. 2017) analysis have a standard prehodiernal past meaning and a future meaning, and have predicative force (as is the case in Sanapaná).

[18] Examples (60a-c) were provided by Rodrigo Villagra Carrón, an anthropologist who wrote his doctoral dissertation with the Angaité. They were accompanied by free translations in English, the glosses and morpheme segmentations were made by me and should thus not be taken as the result of any serious analysis of language-specific Angaité categories but rather as categories loosely corresponding to the Sanapaná and Enenlhet ones discussed here.

Additionally, he argues that they have pragmatic implications in the domain of stance and footing, expressing information both about the source of the information for the speaker, and the expectations the speaker has about the prior knowledge of his interlocutor about the subject of the conversation. Furthermore, it is possible in Enxet for the three particles to occur in clause-initial position (i.e. preceding the verb), in which case they also contribute semantics of negation (62).

(62) Enxet (Enlhet-Enenlhet, Elliott p.c. 2017)
ałta ejke e-taqna-wejkha łama enłoho mʔa ap-jáp
PST CF 2/3M-twist-REP one IRR that 2/3M-father
'His father had never at any time displeased him.'

However interesting and worthy of further (descriptive and comparative) investigation these findings are, there are not enough systematic data available for me to take up either Angaité or Enxet in the historical comparison in the next section.

6.1.2 Enlhet

For Enlhet, more systematic data are available, all taken from Kalisch (2009).[19] From these data, it can be gleamed that Enlhet has cognates of all three Sanapaná particles discussed in this thesis. For *ałta*, the Enlhet cognate seems to generally have the same form as the Sanapaná version. Furthermore, its most typical position also seems to be in the second clausal position following the main verb, in which case it temporally locates the state of affairs expressed by this verb in the past (63) before today.

(63) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 123)
e-l-pejwe:s-a-mk-eʔ ałta enłet majkaʔa ałta kwesejʔ
1SG.O-DISTR-greet-FACT-CPL-PRIM PHOD man/enlhet visitor PHOD because
'The people greeted me, because I was a visitor.'

Just as in Sanapaná, *ałta* combines with a verb in the realis/factive form (*elpejwe:samkeʔ* 'they greet me') in order to convey past temporal reference. Furthermore, it is shown here that *ałta* can also be combined with a noun (*majkaʔa* 'visitor'). From this example, and according to Kalisch' (2009: 119-20) exposition, it seems clear that this noun is used as an independent predicate (i.e. meaning 'to be a visitor'). Since, again according to Kalisch (2009: 122), Enlhet nouns can be used as independent predicates without being followed by a particle, the function of *ałta* here is to make this predicativity explicit and locate it in the past.

Apart from locating the time at which a quality is true about a certain referent in the past (for example, the time when the 'I' possessed the quality of being a visitor), *ałta* in combination with a noun can also receive a discourse coherence function. It can locate in the past the time at which a referent was first activated in the conversation (64a). When combined with the reportative particle *neʔ*, it indicates that the referent was talked about in the past (64b). In sum, the functions of *ałta* in Enlhet seem very similar to those in Sanapaná. Even though there seem to be a number of differences (for example, *ałta* was never found locating the modal load of the irrealis auxiliary in the past), these might upon investigation of more data, prove to not be differences at all.

[19] The glosses are again taken over and adapted to the conventions of this thesis; Spanish free translations are translated into English by me.

(64) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 134; 125)

a. *k-etsep-kek* *nek* *semheŋ* *ałta* *k-jew-eyʔ*
2/3M-die-FACT.PRIM RPT dog(F) PST 2/3F-become_large-INF
'It is said that the large dog which we saw before died.'

b. *aŋ-w-aʔae-kt-aa* *nło* *ŋ-ja:łengʔ* *neʔ* *ałta*
F-arrive:LOC- PST-VENT-PRIM INT F-younger_sibling RPT PST

ŋ-kelʔapa *nak* *ła*
F-old_lady AFF DEM.PROX
'Did the sister (we talked about her before) of this old lady come?'

Phonetically, it seems that *ałta* is more stable in Enlhet than in Sanapaná: the first syllable is maintained even when it follows a vowel (as in *majkaʔa ałta* in 63), and the lateral fricative never seems to be replaced with a glottal stop. The only variation seen is that both *a*'s may become more closed in certain phonetic contexts. When *ałta* follows a word-final *e* or *eʔ*, the first *a* can become an *e* (as in *ŋja:łengʔ neʔ ełta*, 'her younger sister', Kalisch 2009: 125); whereas when it precedes a word-initial vowel, the final *a* may also turn into an *e*, in the process deleting this word-initial vowel with which it collides (as in *sekaʔa ałte ŋ-naʔatkeʔ*, 'the three children'). These two examples are the only ones given, however, which makes it impossible to accurately describe the phonetic contexts that trigger these changes.

A last observation to be made is that the noun for 'night', which is *pesa:sep* in Sanapaná, is *ałtaʔa* in Enlhet. Since it is common in the Enlhet-Enenlhet languages for bare nouns to be used as time adverbials (in which case *ałtaʔa* would have most likely meant 'last night'), it is not inconceivable that this noun might have grammaticalized into a particle expressing past temporal reference. More diachronic comparative research into the Enlhet-Enenlhet languages is necessary to confirm or falsify this hypothesis, however.

The particle *łke*, then, does have a different form in Enlhet than in Sanapaná. In Kalisch (2009), the allomorphs *łka:k*, *ła:k*, and *ła:* are found. The low central vowel here seems similar to the one in Angaité, however. When a form of this particle follows the main verb (65a), it always seems to locate the event expressed by this verb in the hodiernal past, never in the immediate future, pointing to a functional difference with the Sanapaná particle as well.

(65) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 113; 130)

a. *ak-paʔmehet-kas-kek* *ła:k* *ta:ta*
1SG-speak-CAUS-PRIM.FACT HOD father
'I (just) spoke with my father'

b. *ka-m-a* *ła* *//* *sawow* *ła:* *k-jew-ejʔ*
2/3F-grab/buy-IMP DEM.PROX PAUSE knife(F) HOD 2/3F-become_big-INF
'Grab the big knife that we talked about!'

When combined with a noun (as in 65b with *sawow* 'knife' and in 66a with *nsekaʔa* 'child'), this particle is often used with a discourse coherence function as well, indicating that the referent was present in the conversation (65b) or in the environment of the participants in the conversation (66a) recently. Even though Kalisch (2009) does not elaborate on this topic, it can be imagined that *ła:k* marks referents that were discussed more recently than *ałta*. No examples are found of this particle being used in a way analogous to *ałta* in 63a, which it locates the time at which a certain quality is true for the referent of the noun in the past. Once again, however, this might merely be a gap in the data: Kalisch (2009: 132) argues that all temporal particles can indicate "the time of existence of a referent, the time of a relation between referents, or the time of a previous reference to a referent".

Phonetically, a number of rules can be established according to Kalisch (2009: 114) which govern the form of this particle. Whether the final *k* appears or not depends on the initial sound of the following word. If this is the *a* of a person-marking prefix, both this *a* and the final *k* of the particle are typically deleted, as in the case of *ła: k-jew-ejʔ* in (65b), instead of the expexted *ła:k ak-jew-ejʔ*. Whether the first *k* of the particle appears or not depends on the final sound of the previous word: if this is a vowel, the *k* appears (as in 66a), if this is a consonant, the *k* does not appear (as in 65b above).

(66) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 121; 137)

a. *aŋ-law-ke* *jla* *n-sekaʔa* *łka:k*
2/3F-get_angry-FACT.PRIM INFR F-child HOD
'Apparently, the child that was in our focus before got angry.'

b. *v-aʔa-kt-a:k* *mokham* *neŋ-el-anj-a-łk-o* *łka:k*
arrive.LOC-FACT-VEN-PRIM again 1PL-DISTR-advise-FACT-PAS-SBJV HOD

aŋ-aʔh-ek *tataʔa* *pak*
1PL-kill-POT[20] chicken POSS.M
'The one whose chicken they asked us to kill arrived again.'

Logically, this should result in four possibilities: *łka:k* when following a vowel and not preceding a personal prefix, *łka:* when following a vowel and preceding a personal prefix, *ła:k* when following a consonant and not preceding a personal prefix, and *ła:* when following a vowel and preceding a personal prefix. Nevertheless, *łka:* is unattested: in (66b), the particle follows a vowel and precedes a personal prefix, but still we see the form *łka:k angaʔhek* rather than the expected *łka: ngaʔhek*. This indicates either that the rule of final *k*-deletion before a vowel is optional, or that it does not apply in contexts where insertion of the first *k* is triggered by a preceding vowel.

The last particle, then, takes the form *sa:t(e)* in Enlhet (similar to the Enxet version, as shown in subsection 6.1.1). It can follow both an irrealis-marked verb form, which it then temporally locates in the future (67a), or an imperative (67b), in which case no information is given about the exact semantic contribution of *sa:t*. The particle can also follow a noun, where it then fulfils one of the three functions of temporal particles listed above (i.e. it locates the moment at which a referent existed, the time of a relation between referents – e.g. a possessive relationship – or the time of a previous appearance of the referent in the discourse or the context). In the case of (67c), it locates the existence of the referent of *ałtaʔa* in the future (i.e. it indicates that the specific night about which the speaker talks still has to come). Although Kalisch (2009) does not argue this specifically for the particle *sa:t*, he does explain that all the temporal particles make the inherent predicativity of the noun they follow explicit. As such, a direct translation of (67c) would be 'Maybe it will rain. It will be tonight.'

(67) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 133; 129; 144)

a. *ma:me-k-ʔa: ʔ* *//* *e-l-je:s-ek* *sa:t*
rain-FACT-PRIM PAUSE 2/3M-DISTR-get_wet-IRR FUT

enłet *apk-elejw-a* *amjep*
person 2/3M-go-SBJV field
'It is raining. Those who are in the field will get wet.'

b. *eŋj-a:n-e* *sa:t* *e-tne-hek* *maʔa*
2/3M-advise-IMP FUT 2/3M-do-IRR DEM.DIST
'Advise him to do it (that which we talked about before).'

[20] Kalisch' (2009) *potential* corresponds to my *irrealis* in Sanapaná.

c. *ka-mama:-k* *la* *ałtaʔa* *nak* *sa:t* *hajʔ*
2/3F-rain-POT POT night AFF FUT DEM.PROX
'Maybe it will rain tonight.'

Ultimately, this particle as well seems to have two allomorphs. In (68) below, the form *sa:te* appears when the particle is followed by the demonstrative *mʔa*.

(68) Enlhet (Enlhet-Enenlhet, Kalisch 2009: 132)
a-łong *sa:te* *mʔa* //
1SG-travel.POT FUT DEM.DIST PAUSE

ap-w-aʔa-kt-amo *sa:t* *ta:ta*
2/3M-arrive.LOC-FACT-VEN-SBJV FUT father
'I am going to travel (there) when my father has arrived.'

The explanation that first comes to mind is that *sa:te* is the full form, and that the final *e* is deleted when the particle precedes a word starting in an *e* such as *enłet* or *etnehek* in (67a-b). Nevertheless, examples such as (67c) and the second part of (68), where the *e* is deleted before an *h* and before a *t* contradict this hypothesis. One further phonetic particularity of this particle that must be noticed is that, as opposed to Sanapaná *hata*, it never seems to delete its first syllable. More data is needed, however, in order to accurately describe the conditioning of these allomorphs.

6.1.3 Guaná

Guaná as well shows cognates of the three temporal particles under discussion in this thesis, the full forms of which are *ałta*, *łkeʔek* and *hata* according to Unruh & Kalisch (1999).[21] The use of *ałta* seems straightforwardly similar to that of its Sanapaná and Enlhet counterparts. In (69a) it is used in the second position of the clause, immediately following the main verb. As such, it has prehodiernal past reference function. In (69b), it can be seen that *ałta* can follow the clausal negator *maʔ*, in which case it negates the occurrence of the event expressed by the verb (in this case, searching fish) at a certain point in the past.

(69) Guaná (Enlhet-Enenlhet, Unruh & Kalisch 1999: 22; 58; 36)
a. *as-wet-ʔaʔ* *ałta* *ała*
1SG-see-FACT.PRIM PHOD wax_palm
'I saw a wax palm tree'

b. *maʔ* *ałta* *e-tenjaʔ-ak* *kelasma* *tata*
NEG PHOD 2/3M-search-SEC[22] fish father
'My father did not search for fish (at a certain point before today).'

c. *aŋ-kanakh-e* *ełta* *lheja* *jaqtepa* *aknem* *ałta*
2/3F-plant.FACT-PRIM PHOD PRO:2/3F squash sun/day PHOD
'You planted squash yesterday.'

[21] The format of Unruh & Kalisch (1999), a learner's grammar of Guaná, provides Guaná sentences with free translations into Spanish, and a number of comments on the grammatical structure of the language. There are, however no systematic glosses. The glosses given here are constructed by me based on the exposition in the book and my own knowledge of Sanapaná and are therefore provisional. Any errors are my own.

[22] The secundive is, according to Unruh & Kalisch (1999) and Unruh et al. (2003), one of the modes of subordination in the Enlhet-Enenlhet, automatically used when the verb does not occur in its unmarked clause-initial position.

Only one example is given where *ałta* follows a noun (69c). It is not clear whether in this language the syntactic sequence *aknem ałta* has lexicalised as strongly as it has in Sanapaná, but in any case this example provides evidence that *ałta* can locate the existence of referents of nouns in the prehodiernal past. Whether it also serves to make the predicativity of nouns overt cannot be determined from these data.

From (69c), the allomorphy affecting Guaná *ałta* can also be seen: just as in Enlhet, when it follows a word-final *e*, the first *a* of the particle can become more close and convert into an *e* as well (Unruh & Kalisch 1999: 43).

Secondly, *łkeʔek* can also be used in combination with a verb, expressing a hodiernal past (70a). No examples are found, however, of it expressing an immediate future or a prospective aspect, and neither do Unruh & Kalisch (1999) allude to any such use at any point in their exposition. It thus seems that the use of this particle in Guaná lies closer to that of its Enlhet cognate than to that of the Sanapaná one. This particle can also follow the future/desiderative auxiliary (70b), in which case it locates the desire expressed by this auxiliary in the hodiernal past (and not necessarily the event expressed by the main verb); and just as *ałta*, it can follow the negation in order to locate the non-occurrence of the event in the hodiernal past (70c).

(70) Guaná (Enlhet-Enenlhet, Unruh & Kalisch 1999: 64; 26; 58)

a. *nen-tavaŋ-ke* *łkeʔek* *ja:tekjełna*
1PL-eat.FACT-PRIM HOD orange
'We just ate oranges.'

b. *as-maka* *łkeʔek* *aŋ-ven-ekh-ak* *sa:nja* *paktek*
1SG-want HOD 1SG-plant-REP-FUT watermelon seed
'I just wanted to plant watermelon seeds.'

c. *ma* *łkeʔek* *a-tenjaʔ-ak* *jentapa*
NEG HOD 1SG-search-SEC wood
'I did not just search for wood.'

When co-occurring with a noun, several interpretations are possible. In (71a), it seems that *łkeʔek* locates the referent of the noun *saasekhe* 'morning' in the hodiernal past, yielding a reading 'this morning'. In (71b), on the other hand, it seems as though *łkeʔek* still has clausal scope, even though it follows the noun *nałma* 'forest' rather than the verb *sjejha:maha* 'walk'. From the example sentences, it is also not entirely clear whether the use of the particle here implies that the noun is used predicatively.

(71) Guaná (Enlhet-Enenlhet, Unruh & Kalisch 1999: 54)

a. *akhawe'* *sa:sekhe* *łkeʔe* *p-łej-ak*
no morning HOD 2/3M-travel-SEC
'No, he went today in the morning.'

b. *nałma* *łkeʔe* *s-jejha:m-aha*
forest HOD 1SG-walk.FACT-SEC
'Yes, I just walked in the forest'

Unruh & Kalisch' (1999) data show evidence for a morphophonological process directly affecting *łkeʔek*. When immediately followed by the *a* of a person-marking prefix, (71a), both this *a* and the final *k* of *łkeʔek* can be elided. Furthermore, this particle affects the preceding word: it can trigger the deletion of a word-final *k* or glottal stop (as is the case in 70a, where the verb form in isolation would be *nentavangkeʔ* according to Unruh & Kalisch 1999: 71), and it can trigger the shortening of word-final suffixes of the preceding word (causing amongst others *-aha* to become *-a*, Unruh & Kalisch 1999: 35).

The last particle, *hata*, can just as in Enlhet follow an irrealis/future-marked verb (72a), conveying future temporal reference. The particle in this case indicates that the event will take place in a more distant future than an event expressed in a clause with the future auxiliary (Unruh & Kalisch 1999: 35). When *hata* follows an imperative (72b), it portrays the order as less urgent and less immediate (Unruh & Kalisch 1999: 35). Unruh & Kalisch (1999) do not give any examples of *hata* following a noun. It is thus impossible to know, based on this data, whether *hata* has the same function as a marker of predicativity in Guaná that it does in the other languages of the family.

(72) Guaná (Unruh & Kalisch 1999: 42; 50)

a. *a-to-k* *hata* *s-to-ma*
1SG-eat-FUT FUT 1SG-eat-INF
'I will eat my food.'

b. *ko-taw* *hata* *k-to-ma*
2/3F-eat.IMP FUT 2/3F-eat-INF
'Eat your food!'

c. *a-mo* *ta*
1SG-buy FUT
'I am going to buy it.'

d. *a-jejhama* *ta*
1SG-walk.FUT FUT
'I am going to go for a walk.'

Phonetically, the first syllable of *hata* is lost when the particle immediately follows a vowel. This can also be the case when it follows a future verb form ending in *k*, in which case this word-final *k* is elided as well (as in 72c, where the full form of the verb is *amok*). Another way in which *hata* can influence the form of the preceding word, is that just like *łkeʔek*, it can shorten wordfinal *-aha* to *-a* (as in 69d, where the full form would be *ajejhamaha*).

6.1.4 Enenlhet

The last language of the Enlhet-Enenlhet family, Enenlhet, also has cognates of the three particles under scrutiny in this paper. The Enenlhet cognate of *ałta* has the same base form, and is used in the same base context: following a verb in order to temporally locate it in the prehodiernal past (73a). As in Guaná, it can also immediately follow the clause-initial negation *ma*, in which case it locates the non-occurrence of the event expressed by the verb in the past before yesterday (73b). (73c) shows that *ałta* can follow the interrogative particle *se* in which case it indicates that the question asks about an event that happened in the past. Additionally, there seems to be a possibility that is unattested in any of the other Enlhet-Enenlhet languages: *ałta* can appear clause-initially without any material preceding it, without giving the clause a negative reading (73d).

(73) Enenlhet (Unruh et al. 2003: 60; 92; 194; 242)

a. *ap-łeng-kek* *ałta* *Juan* *ap-jataʔ*
2/3M-travel.FACT-PRIM PHOD Juan 2/3M-grandfather
'Juan's grandfather went away (before today).'

b. *ma* *łta* *ŋko-ten-ak* *meme* *pełasep* *ałta*
NEG PHOD 2/3F-sleep-SEC mother night PHOD
'Mother did not sleep the other night.'

c. *¿se łta k-wet-ʔo knem ałta je? loma-neten ałta*
INT PHOD 2/3F-see-SBJV sun/day PHOD INT airplane PHOD
'What did you see yesterday? It was an airplane.'

d. *ałta s-ten-akh-e ŋkeʔe ak-wa-j-ʔo nak menasma*
PHOD 1SG-sleep-REP-PRIM DEM.PROX 2/3F-come-STX-SBJV PRES snake
'This is the place where I slept when the snake came.'

When used with nouns, Unruh et al. (2003: 194) give clear examples of *ałta* being used to specify the predicativity of a noun: *loma-neten ałta* in (73c) does not mean 'something that was a plane' or 'the plane which we talked about,' but rather 'it was a plane'. Furthermore, it can again be used to locate the existence of the referent of that noun in the prehodiernal past (as in 73b, where *pełasep ałta* refers to 'a night before today'), or to indicate that this referent was present in the immediate environment of the speaker and/or listener in the past (74a). Alternatively, when *ałta* follows a sentence-initial noun it can seemingly have clausal scope (74b), locating the whole event of 'seeing a dog' in the past rather than the previous presence of the dog.

(74) Enenlhet (Unruh et al. 2003: 74; 28)
a. *as-weta-kh-ek łek peletaw ałta*
1SG-see-REP-PRIM HOD knife PHOD
'I saw/found the knife again that was there before.'

b. *semheŋ ałta s-wet-ʔak*
dog PHOD 1SG-see-SEC
'I saw a dog.'

In this language as well, lastly, *ałta* can take part in a number of morphophonological processes. In (73b-c, *ma łta* and *se łta*), it can be seen that the initial *a* is lost when the particle follows a word-final vowel. Additionally, *ałta* can trigger the elision of final *k* of the preceding word (75), in which case the initial *a* of the particle is lost as well. This does, however, not seem as pervasive as it is in Sanapaná, for example: examples such as (73a) where the final *k* and initial *a* stay intact are plentiful.

(75) Enenlhet (Unruh et al. 2003: 60)
n-łeŋ-ke łta meme aknem ałta
2/3F-travel-PRIM PHOD mother sun/day PHOD
'Mother travelled yesterday.'

The cognate of Sanapaná *łke* has again a number of allomorphs in Enenlhet. The initial *k* only appears when the particle follows a vowel (or a *k* which it deletes), and the final *k* can also be elided when the following word starts with a vowel, resulting in four variants (*łkek, łke, łek* and *łe*). Again, the process of final *k*-deletion is not as pervasive in Enenlhet as it is in Sanapaná since examples like (76a) are often found (where one would expect *apteŋke łkek* if the aforementioned rule were to apply). Since these processes are the same as the ones we saw before, I will not separately exemplify all of them here. One extra form of the particle is found that we did not encounter in any other language: when *łkek* occurs sentence finally, it takes the form *ł(k)a jek* (76a), without this having any effects on the semantics of the particle (Unruh et al. 2003: 58).

When following a verb, it once again locates the event to which this verb refers in the hodiernal past (76a-c), and when following interrogative particle *anło* (76d), it indicates that the speaker is asking about an event in the recent past.

(76) Enenlhet (Unruh et al. 2003: 60; 82; 74; 44).

a. *ap-teŋ-kek* *lhek* *nematka* *ła jek*
2/3M-sleep-PRIM HOD child HOD
'The child that was here just fell asleep.'

b. *makʔak* *nak* *łe* *ŋ-kan-ekh-ak* *meme*
want RPT HOD 2/3F-cook-REP-SEC mother
'Mother said she is going to cook.'

c. *as-makʔak* *łek* *a-kne-kh-ak* *samaneja*
1SG-want HOD 1SG-plant-REP-FUT watermelon
'I just wanted to plant watermelon.'

d. *¿ap-łeŋ-ke* *nło* *łkek* *tata?*
2/3M-travel-PRIM INT HOD father
'Did father leave?'

Interestingly, *łkek* can be used in clauses like (76b) where it follows the reportative particle *nak*. In this case, the hodiernal past semantics seem to be associated with the reportative only, rather than with the clause as a whole: the *saying* event took place earlier on the day of this utterance, whereas no information is given concerning when the cooking took place, if it did at all. The future meaning here is entirely expressed by the modal *makʔak*, so it cannot be said that *łkek* in these examples has a prospective aspect meaning in the way that it does in Sanapaná. If the reportative particle is absent (as in 76c), *łkek* locates the desire or intention expressed by the auxiliary in the hodiernal past.

When used with nouns, *łkek* can seemingly also be used with clausal scope, as exemplified in (77a), with discourse coherence function (77b), to refer to an entity that was recently present in the extralinguistic context (76a), or to locate the existence of a referent in the hodiernal past (77c). Additionally (77d) is a clear example of *łkek* being used to overtly express the predicativity of a noun.

(77) Enenlhet (Unruh et al. 2003: 28; 56; 68; 6)

a. *menasma* *łke* *s-wet-ʔak*
snake HOD 1SG-see-SEC
'I (just) saw a snake.'

b. *ŋ-van-ekh-a* *tajke* *jaqtepa* *łkek*
1SG-cook-REP-FUT later gourd HOD
'Later I will cook the gourds that we just talked about.'

c. *ap-sant-anj-ek* *łek* *seta* *tape* *sosekhek* *łek*
2/3M-bring-DIR-PRIM HOD grandfather chicken morning HOD
'This morning, my grandfather brought a chicken.'

d. *¿nohajkok* *anło* *n-san-t-ak* *jaja?* — *maʔŋ-vengka* *łka jek.*
jar INT 2/3F-bring-SEC sister no 2/3F-pot HOD
'Is it a jar that my sister brought? – No, it was a cooking pot.'

The last Enenlhet particle, *hata*, has a straightforward future temporal reference meaning in most of its uses. In sentences such as (78a), hata follows a future/irrealis-marked main verb, in which case it locates the event expressed by it in the future, whereas in (78b), it follows an imperative, imbuing the order with a weakened sense of urgency (as it did in Guaná). Examples of *hata* occurring after verbs that are preceded by the auxiliary *makʔak* are unattested in Unruh et al. (2003). According to these authors (2003: 52), this has to do with a semantic constraint: *makʔak* presents an event as about to happen, whereas *hata* presents it as a more remote future.

(78) Enenlhet (Unruh et al. 2003: 24; 32)

a. *a-ɬ-ek* *hata*
1SG-sit-FUT FUT
'I am going to sit down.'

b. *e-sant-a* *ta* *p-taɬnema*
2/3M-bring-IMP FUT 2/3M-shirt
'Bring your shirt!'

As seen in (79a), Enenlhet allows *hata* to appear clause-initially, without contributing negative semantics to the clause.

(79) Enenlhet (Unruh et al. 2003: 202; 36; 74)

a. *hata* *a-m-o* *k-mopajʔa* *nak*
FUT 1SG-buy-FUT 2/3F-white PRES
'I am going to buy the white one.'

b. *a-te ŋj-ak* *hata* *nen-to-ma*
1SG-search-FUT FUT IMPERS-eat-INF
'I am going to search for food.'

c. *a-teŋj-a* *ta* *vakkahak*
1SG-search-FUT FUT notebook
'I am going to search for a notebook.'

d. *ŋko-mma-k* *hate* *ɬaha*
2/3F-rain-FUT FUT maybe
'Maybe it will rain.'

Example (78b) also shows the variation undergone by this particle when it follows a word-final vowel: the first syllable *ha-* is lost in these cases, and also optionally when it follows a *k*, which in turn is deleted as well (as shown in the pair 79b-c). When *hata* precedes a word-initial *(k)e*, the final *a* of *hata* can assimilate to an *e*, and this word-initial *e* itself is deleted (as in 79d where it precedes *keɬaha* 'maybe').

When used with a noun, *hata* can have scope over the whole clause, as in (80a) where the sentence is about a future searching event, not an entity that will be a fruit in the future. The most frequent interpretation of *hata*, however, is that it locates the referent's existence in the future, since the discourse coherence function seems quite incompatible with the future semantics of *hata*. This is the case in (80b). This example merits special attention, since it forms a direct contrast with the Sanapaná example in (49b), repeated here for convenience as (81). In the two cases, the noun for 'moment' is followed by a particle to express an immediate future. In Sanapaná, speakers will generally use *ɬke* in such a context whereas in Enenlhet, *hata* is apparently the default choice. It thus seems that *ɬke* has not been co-opted to express prospective aspect in Enenlhet in the same way that it has been in Sanapaná.

(80) Enenlhet (Unruh et al. 2003: 48)

a. *jajet* *hata* *a-teŋj-ak* *jenaqteso*
jujube_tree FUT 1SG-search-FUT afternoon
'I am going to look for jujube fruits this afternoon.'

b. *koɬawek* *hata* *a-kn-ekh-ak* *samaneja* *paktek*
moment FUT 1SG-plant-REP-FUT watermelon seed
'I am going to plant watermelon seeds right now.'

(81) Sanapaná (FN1.21)
keɬwoje=ɬke *moʔo* *o-w-ata*
moment=IMM 1.AUX.IRR 1SG-arrive-VENT.IRR
'I will soon arrive (back) here'

6.2 Implications for Enlhet-Enenlhet genetic relations

In the previous section, I presented evidence that the Sanapaná particles *ałta*, *łke* and *hata* have cognates in all its fellow Enlhet-Enenlhet languages. The presence of these particles across the language family implies that they are a feature which was inherited from a common ancestor, rather than an innovation. In the following discussion of this hypothesis, and the application of the historical glottometric approach to these data, Enxet and Angaité will be disregarded because of the lack of available data. The particles take the forms summarised in table 5 below.

Sanapaná	Enenlhet	Guaná	Enlhet	(Enxet)	(Angaité)	*proto-Enlhet-Enenlhet
łkeʔek *ʔkeʔek* *łkeʔe* *łke* *ʔke*	*łkek* *łek* *łke* *łe* *ła iek*	*łkeʔek* *łkeʔe*	*łka:k* *ła:k* *ła:*	*(łek)* *(łejk)*	*(łajk)* *(łkek)*	**łkeʔek*
ałta *łta* *ʔta*	*ałta* *łta*	*ałta* *ełta*	*ałta* *ełta* *ałte*	*(ałta)*	*(ałta)*	**ałta*
hata *ta*	*hata* *ta* *hate*	*hata* *ta*	*sa:te* *sa:t*	*(sate)* *(sat)*	*(hata)*	**sa:te*

Table 5: Attested forms of the temporal particles across the Enlhet-Enenlhet languages

As shown in table 5 and demonstrated in the previous section, the full forms of these particles are very similar across all the languages in which they are attested. With regards to the hodiernal past particle, the consonantal frame can be reconstructed rather easily: *łkVʔVk*. All languages have forms initiating in the lateral fricative and ending in *k*, and five of them have forms with the initial *k*. I also assume a *VʔV* sequence rather than a long vowel, since it seems to be a tendency in the Enlhet-Enenlet languages to lose intervocalic glottal stops. Which vowel must be reconstructed, is less easy to determine, however. I hypothesise tentatively that this would have been an *e* rather than an *a*. If languages such as Sanapaná and Guaná innovated the original *aʔa* sequence of this particle to *eʔe*, one would expect to find other examples of such a sound change. The *aʔa* cluster is well attested, however, in both these languages (in words such as *majkaʔa* 'visitor'). For these reasons, I posit **łkeʔek* as the protoform for this morpheme.

For the second particle, the prehodiernal past, there are far fewer difficulties: all six languages have an attested form *ałta*, which adapts in different ways to its phonetic environment. This form, therefore, can safely be reconstructed for the protolanguage.

The last particle, ultimately, is reconstructed here as *sa:te*. The initial consonant is an *s* in two of the languages, and an *h* in the four others. Although the *h*-form is attested in more languages, I still opt for the *s*-form in my reconstruction. Since the *s*-form occurs in the two languages of Unruh & Kalisch' (2003) Western Enlhet-Enenlhet branch, and the *h*-form in the Eastern Enlhet-Enenlhet languages, it is plausible that the crosslinguistically common *s* > *h* simply took place in proto-Eastern Enlhet-Enenlhet, and was then passed down into its daughter languages. For the first vowel, I reconstruct a long one, because of the aforementioned tendency of the Enlhet-Enenlhet languages to shorten long vowels, making *a:* > *a* a more plausible sound change than *a* > *a:*. The *t* segment occurs in all languages and is thus uncontroversial. For the last vowel, I reconstruct an *e*. In the previous paragraph, I posited an *a* > *e* sound change in the Sanapaná, Enlhet and Guaná forms of the hodiernal past participle, whereas the

Enlhet form maintained the *a*. It would therefore be rather inconsistent to now posit a protoform *sa:ta* which would have undergone this *a* > *e* sound change in Enlhet but not in any of the other languages. The cause of the *e* > *a* sound change in Sanapaná, Enenlhet, Guaná and Angaité might then possibly have been some form of vowel harmony in proto-Eastern Enlhet-Enenlhet.

Based on the full forms of these particles, the following genealogical hypothesis in terms of innovations can be established (as summarised in table 6, where the bold print indicates which languages group together with respect to a certain feature): Sanapaná and Guaná seem to be closest together. Enlhet differs from them and Enenlhet in the vowels of the prehodiernal particle (where Enlhet underwent an *e* > *a* sound change) and the future particle (where the three other languages lost the long first vowel and underwent an *a* > *e* sound change. Furthermore, an *s* > *h* innovation took place in this same future particle in Sanapaná, Guaná, Enenlhet and (apparently) Angaité. Enenlhet also shows an innovation, in that it is the only language which lost the long vowel in the hodiernal past particle. Therefore, based purely on these full forms, one can hypothesise that the original proto-Enlhet-Enenlhet dialect continuum took the following form:

Enlhet – Sanapaná – Guaná – Enenlhet

This hypothesis corresponds reasonably well to that proposed by Unruh & Kalisch (2003) and discussed in the introduction of this thesis.

Enlhet	Sanapaná	Guaná	Enenlhet
Low vowel in łka:k	**Mid vowel in łkeʔek**	**Mid vowel in łkeʔek**	**Mid vowel in łkek**
s-form of sa:te	**h-form hata**	**h-form hata**	**h-form hata**
Long vowel in sa:te	**Short vowel in hata**	**Short vowel in hata**	**Short vowel in hata**
Mid vowel in sa:te	**Low vowel in hata**	**Low vowel in hata**	**Low vowel in hata**
Long vowel in łka:k	**Long vowel in łkeʔek**	**Long vowel in łkeʔek**	Short vowel in łkek

Table 6: Distribution of the features of the full forms of the particles in the Enlhet-Enenlhet languages

In terms of the morphophonological processes in which these particles take part, it can be seen that the loss of both the initial *k* when *łkek* follows a consonant, and the final *k* when it precedes a personal prefix, are common processes of Enenlhet and Enlhet, whereas they seem absent from Sanapaná and Guaná. Sanapaná seems to be the only language where the initial lateral fricative can be glottalized, and Enenlhet appears to be the only family member which has a special clause-final variant of the particle. For *ałta*, two patterns of variation can be seen. In Sanapaná and Guaná, the initial vowel is deleted when the particle follows a vowel, which is not the case in Enenlhet and Enlhet. In the latter two languages, however, this initial vowel can assimilate to a preceding *e*, which is unattested in Sanapaná and Guaná. Again, Sanapaná is the only language which shows optional glottalization of the lateral fricative. For *hata*, ultimately, Enlhet is the only language where the initial syllable consistently remains stable, whereas this can be elided when following a vowel or *k* in Sanapaná, Enenlhet, and Guaná. Enlhet is also the only language where the final vowel can be deleted. These morphophonological processes are summarised below in table 7 where, again, the bold print indicates which languages pattern together regarding a certain feature.

As can be seen, the morphophonological features associated with these particles show a slightly different distribution than the full forms of the morphemes: they correlate best with the following proto-Enlhet-Enenlhet dialect continuum:

Enlhet – Enenlhet – Guaná – Sanapaná

Within this continuum, Guaná and Sanapaná share five of the six relevant features and are thus most closely related. Enenlhet shares two features with these two languages, and an additional one with Guaná alone. Therefore, it seems to occupy an intermediate position between the Guaná-Sanapaná cluster and Enlhet. Enlhet, ultimately, shares three features with Enenlhet, two with Guaná and only one with Sanapaná, locating it on the other side of Enenlhet than Guaná and Sanapaná.

Enlhet	Enenlhet	Guaná	Sanapaná
Final vowel of sa:te can fbe lost	**Final vowel of hata remains stable**	**Final vowel of hata remains stable**	**Final vowel of hata remains stable**
Initial syllable of sa:te remains stable	**Initial syllable can elide after vowel**	**Initial syllable can elide after vowel**	**Initial syllable can elide after vowel**
Initial vowel of ałta assimilates after e	Initial vowel of ałta assimilates after e	**Initial vowel of ałta lost after vowel**	**Initial vowel of ałta lost after vowel**
łka:k loses initial k after consonant	łkek loses initial k after consonant	**łkeʔek keeps initial k after consonant**	**łkeʔek keeps initial k after consonant**
No special clause-final variant	łkek becomes ła iek clause-finally	**No special clause-final variant**	**No special clause-final variant**
No glottalization of ł	No glottalization of ł	No glottalization of ł	**ł of łke and ałta can be glottalised**

Table 7: Phonological processes undergone by the Enlhet-Enenlhet particles

In terms of the functions fulfilled by these particles, it is harder to come to a clear conclusion about which languages are most closely related, because of the significant possibility of gaps in the data. The most striking difference between the languages in this area is the function of the *łke* particle: in Enlhet, Enenlhet, and Guaná, this seems to be a hodiernal past marker, locating either the event expressed by the verb it follows or the existence of the referent of the noun it follows earlier on the day of the utterance. In Sanapaná (and, possibly, Enxet) it can function as an immediate aspect marker, locating the existence of the event expressed by the main verb (if it follows the auxiliary) in the immediate future.

The data presented in this thesis point towards a number of further possible functional differences between the use of the particles in these languages. For example, the use of *łkeʔek* following a noun but having scope over the clause as a whole is only encountered in Enenlhet and Guaná, not in Enlhet and Sanapaná. Nevertheless, the available data are too limited to confidently make this claim, as the lack of such examples in these two languages might prove to be due to gaps in the data. Therefore, I will not consider these features when making my conclusions, leaving their examination for a later study. Table 8 below shows that Sanapaná contrasts with Enlhet, Enenlhet and Guaná regarding the only two functional characteristics of the particles that are taken into account here.

Enlhet	Enenlhet	Guaná	Sanapaná
łka:k locates N in hodiernal past	łkek locates N in hodiernal past	łkeʔek locates N in hodiernal past	**łke can locate N in immediate future**
łka:k locates V in hodiernal past	łkek locates V in hodiernal past	łkeʔek locates V in hodiernal past	**łke can locate V in immediate future**

Table 8: Functional characteristics of the particles in the Enlhet-Enenlhet languages

Altogether, the formal and functional differences between the cognates of *łke*, *ałta* and *hata* point towards a relatively clear proto-dialect continuum from which the present-day Enlhet-Enenlhet languages could have descended. Table 9 summarises all relevant characteristics which have been treated in this section, and indicates (by means of the boldface-regular print contrast) which languages group together with respect to each feature.

	Enlhet	Enenlhet	Guaná	Sanapaná
1	Low vowel in łka:k	**Mid vowel in łkek**	**Mid vowel in łkeʔek**	**Mid vowel in łkeʔek**
2	s-form of sa:te	**h-form hata**	**h-form hata**	**h-form hata**
3	Long vowel in sa:te	**Short vowel in hata**	**Short vowel in hata**	**Short vowel in hata**
4	Mid vowel in sa:te	**Low vowel in hata**	**Low vowel in hata**	**Low vowel in hata**
5	Final vowel of sa:te can be lost	**Final vowel of hata remains stable**	**Final vowel of hata remains stable**	**Final vowel of hata remains stable**
6	Initial syllable of sa:te remains stable	**Initial syllable can elide after vowel**	**Initial syllable can elide after vowel**	**Initial syllable can elide after vowel**
7	**No special clause-final variant**	łkek becomes ła iek clause-finally	**No special clause-final variant**	**No special clause-final variant**
8	**Long vowel in łka:k**	Short vowel in łkek	**Long vowel in łkeʔek**	**Long vowel in łkeʔek**
9	Initial vowel of ałta assimilates after e	Initial vowel of ałta assimilates after e	**Initial vowel of ałta lost after vowel**	**Initial vowel of ałta lost after vowel**
10	łka:k loses initial k after consonant	łkek loses initial k after consonant	**łkeʔek keeps initial k after consonant**	**łkeʔek keeps initial k after consonant**
11	No glottalization of ł	No glottalization of ł	No glottalization of ł	**ł of łkeʔek and ałta can be glottalised**
12	łka:k locates N in hodiernal past	łkek locates N in hodiernal past	łkeʔek locates N in hodiernal past	**łkeʔek can locate N in immediate future**
13	łka:k locates V in hodiernal past	łkek locates V in hodiernal past	łkeʔek locates V in hodiernal past	**łkeʔek can locate V in immediate future**

Table 9: Summary of features relating to temporal predicative particles in the Enlhet-Enenlhet languages

A visual representation of these relations is presented in figure 5 below. From the 4 languages that are considered here, one can form ten subgroups, as shown in table 10 below. As can be seen in both figure 5 and table 10, only five of these groupings are plausible candidates for the position of genetic subgroup of the Enlhet-Enenlhet family based on the data presented in this paper. The triangles and rectangles in the figure group together languages which show the same value for at least one of the features in table 9. How many of the 13 features are shared amongst these languages is indicated by the number in the same colour as the shape. For five of the ten possible subgroups, the languages included in them share none of the features which are discussed here. The Enenlhet-Guaná-Sanapaná subgroup has six shared features, whereas the Enlhet-Enenlhet-Guaná subgroup has three, and the Enlhet-Guaná-Sanapaná, Enlhet-Enenlhet, and Guaná-Sanapaná subgroups all have two.

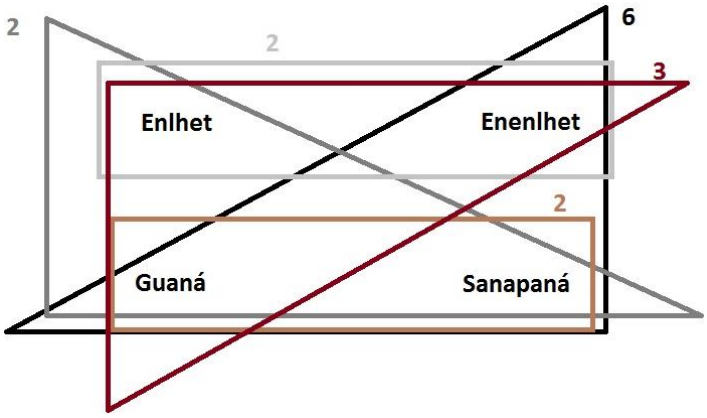

Figure 5: The number of features shared by Enlhet-Enenlhet subgroups

This figure seems, at first sight, to corroborate Unruh & Kalisch' main claim about the genetic relationships between the Enlhet-Enenlhet languages: with respect to the form and function of temporal predicative particles, Sanapaná, Guaná and Enenlhet group together, sharing together six characteristics, which is more than any other grouping of these four languages. As such, Unruh & Kalisch' (2003) Western vs. Eastern Enlhet-Enenlhet distinction seems to be well-founded.

Subgroup	Shared features
Enenlhet-Guaná-Sanapaná	6
Enenlhet-Enlhet-Guaná	3
Enlhet-Guaná-Sanapaná	2
Enenlhet-Enlhet	2
Guaná-Sanapaná	2
Enlhet-Enenlhet-Sanapaná	0
Enenlhet-Guaná	0
Enenlhet-Sanapaná	0
Enlhet-Guaná	0
Enlhet-Sanapaná	0

Table 10: All possible Enlhet-Enenlhet subgroups and the number of features they share

Let us now move on, though, to the question of which of the five subgroups that share any features are supported by the strongest historical glottometric evidence, by calculating the subgroupiness of every one of them as introduced in section 2.3. As a reminder, the formulae for cohesiveness and subgroupiness are repeated in table 11.The Enenlhet-Guaná-Sanapaná subgroup, firstly, is supported by the first six features of table 9 (1-6) and contradicted by the next seven features (7-13). This results in a cohesiveness measure of 6 / (6+7) = 0,46. Multiplied by six, the absolute number of features exclusively shared by these three languages, we get a subgroupiness value *s* of 2,76.

Measure	Formula
Coherence (*c*)	$p / (p + q)$
Subgroupiness (*s*)	$e * c$
p	Number of supporting innovations for a subgroup
q	Number of contradicting innovations for a subgroup
e	Number of exlusively shared innovations for a subgroup

Table 11: Formulae and symbols relevant for the calculation of subgroup strength

The Enlhet-Enenlhet-Guaná subgroup, secondly, is supported by three features (11-13) and contradicted by ten features (1-10), resulting in a cohesiveness value of 3 / (3+10) = 0,23. Since these three languages exclusively share three features (11-13), the subgroupiness of this group is 3 * 0,23 = 0,69. The Enlhet-Guaná-Sanapaná subgroup, then, has two supporting features (7-8) and eleven contradicting ones (1-6 and 9-13). It thus has a *c* of 2 / 2+11 = 0,15, resulting in a subgroupiness value of 0,15 * 2 = 0,30, given that the two supporting features are also exclusive features to this subgroup.

Lastly, there are two relevant subgroups which only contain two members. The Guaná-Sanapaná subgroup, firstly, is supported by ten (1-10) features and contradicted by three features (11-13), resulting in a coherence of 10 / (10+3) = 0,77. Since there are only two exclusively shared featuress, however (9-10), the subgroupiness is still lower than that of Enenlhet-Guaná-Sanapaná: *s* = 2 * 0,77 = 1,54. The last subgroup, Enenlhet-Enlhet, is supported by five features (9-13) and weakened by eight features (1-8), resulting in a coherence of 5 / (5+8) = 0,38 and a subgroupiness of 2 * 0,38 = 0,76, since there are two

exclusively shared features (9-10). All the data concerning the strength of the Enlhet-Enenlhet subgroups are summarised in table 12 below.

Subgroup	Exclusively shared innovations	Coherence	Subgroupiness
Enenlhet-Guaná-Sanapaná	6	46%	2,76
Enlhet-Enenlhet-Guaná	3	23%	0,69
Enlhet-Guaná-Sanapaná	2	15%	0,30
Guaná-Sanapaná	2	77%	1,54
Enlhet-Enenlhet	2	38%	0,76

Table 12: Measures for the strength of each relevant genetic Enlhet-Enenlhet subgroup

As can be seen in this table, the main prediction made by Unruh & Kalisch (2003) is born out by the data on temporal predicative particles presented in this thesis. The best-supported subgroup, with a subgroupiness value of 2,76, is Enenlhet-Guaná-Sanapaná which corresponds to Unruh & Kalisch' (2003) Eastern Enlhet-Enenlhet. Within this branch, however, Unruh & Kalisch (2003) expect Guaná and Enenlhet to cluster together, whereas my data lean towards a stronger grouping of Guaná and Sanapaná: they show a subgroupiness value of 1,54, and a coherence value of 0,77. This indicates that whenever an innovation targeted either Guaná or Sanapaná, there is a 77% chance that it also affected the other one. It was impossible to calculate a subgroupiness value of Enenlhet and Guaná, since the two languages do not exclusively share any innovations with respect to the particles under discussion here.

In sum, this section confirms the intuitions of Unruh & Kalisch (2003) that the Enlhet-Enenlhet language family can better be conceived of as having originated from a dialect continuum with intersecting shared innovations than as a family tree with neat binary branching. Although much research still needs to be done in this field, this study strengthened the evidence for an Eastern-Western divide between Enlhet on the one hand, and Guaná, Sanapaná and Enenlhet on the other hand. This does not mean, however, that Enlhet does not share any features with any of these three languages, as illustrated in table 9. Based on the evidence studied in this thesis, Sanapaná and Guaná seem to group most closely together within the Eastern branch, rather than Enenlhet and Guaná together forming a clear Northeastern Enlhet-Enenlhet subbranch. Preliminary data on Enxet seem to show that this language patterns both with Enlhet (in using the *s*-form of the future particle) and with Enenlhet (in having lost the long vowel in the hodiernal past particle). Angaité seems to be somewhere in the middle of all these languages, since it patterns with Enlhet in the sense that it has an allomorph of the hodiernal past particle with an *a* vowel, it patterns with Enenlhet in that the long vowel in this particle seems to have disappeared, and it patterns with Sanapaná, Guaná and Enenlhet in using the *h*-form of the future particle. On Angaité and Enxet, however, not enough is known in order to meaningfully include them in the methodology applied here.

Chapter 7 Conclusion

After having presented and discussed data from Sanapaná, Enlhet, Guaná and Enenlhet (and a few initial data on Angaité and Enxet) on the form and use of temporal predicative particles, this last chapter will conclude the present thesis. In section 7.1, I summarise the findings presented in the previous chapters, and in section 7.2 I suggest a number of conclusions that can be drawn from this study, and some directions for future research which would be useful with the goal of confirming, refining, or falsifying these findings.

7.1 Summary

This thesis presented both descriptive findings on the Sanapaná language, and historical-comparative findings concerning the whole Enlhet-Enenlhet language family. In chapter 3, I gave an introduction to Sanapaná grammar, focussing on the verbal morphology. In this regard, I argued that the realis verbal paradigm presents a hierarchical alignment system, organised according to a 1 > 2/3 hierarchy. Whenever a first person argument is present in a transitive clause, this argument will be prefixed on the verb (with the *as-* prefix if it functions as an A, with the *e-* prefix if it functions as an O). In clauses without a first person (i.e. 2 => 3, 3 => 2 and 3 => 3 scenarios), the A argument is always prefixed to the verb, with a prefix that distinguishes between masculine and feminine gender. In the irrealis paradigm, no such hierarchical system seems to be present: both the verb-initial auxiliary and the prefix on the verb simply mark the person and number of S and A.

In terms of TAM categories, I argued that Sanapaná does not express tense on the verb. Instead, the categories which are most closely related in their semantics to temporal reference are either modal in core meaning (the irrealis mood, which often has desiderative semantics and is often used with future time reference, and the realis mood, which can be used with both past and present time reference), or aspectual (the imperfective, which is most often used with past semantics but can also express present habits). Another important point which I made in this chapter is that nouns and other parts of speech in Sanapaná can be independently used as predicates, without needing any syntactical or morphological operations.

In chapters 4 and 5, then, the (morpho)phonology, syntactics and semantics of the three particles that are the main focus of this thesis (*ałta*, *łke*, and *hata*) were described. I noted that these are clitics which are syntactically independent but depend phonologically on the verb. When following a verb, *ałta* and *hata* were described as an analytic verbal tense markers expressing, respectively, a situation where the reference time R is the day of the utterance, and the event time E precedes this (E > R(S)), and one

where R is the moment of the utterance and E follows this (R > E). *łke* was described as an analytic verbal aspect marker, expressing immediate phasal aspect: it can portray the event from a viewpoint right before its start, or right after its completion: E2 > R / R > E1.

When combining with nouns, these particles can also receive a temporal interpretation. The only way in which *łke* can fulfil such a function is by locating the previous mention of a referent in discourse (or the previous appearance of this reference in the extralinguistic world). *Ałta* and *hata*, on the other hand, can also temporally locate the existence of a referent, or the time at which a certain quality holds of this referent. The former locates the referent in the past ($t_{NOM} > t_{NP}$ (S)), whereas the latter locates this referent in the future ($t_{NP,}S > t_{NOM}$). Neither of these particles, however, qualifies for the status of a nominal tense marker, according to the criteria of Nordlinger & Sadler (2004a,b) and Tonhauser (2007; 2008).

Additionally, these particles have predicative force: when following a noun, they prompt a predicative reading. Since this predicative reading is also optionally available when the noun is not followed by a particle, these particles specify explicitly the predicativity of the noun, in Kalisch' (2009) terms, rather than making the noun predicative. I furthermore presented my reflections on the status of Sanapaná with respect to the issue of monocategoriality and flexible word classes. At first sight, the ability of nouns to be used predicatively seems to qualify the language for such an analysis. Nevertheless, verbs cannot be used as arguments or referring expressions without being overtly nominalised. For this reason, Sanapaná cannot be seen as a monocategorial language, but rather as an omnipredicative one (all lexemes can predicate, but not all of them can refer).

In chapter 6, ultimately, I broadened the scope of this investigation, and compared my Sanapaná data with the available data on Enlhet, Guaná and Enenlhet, applying to them the method of historical glottometry. I identified thirteen features of the form and function of these particles which show different values in these four languages, and based on these calculated the coherence and subgroupiness of each subgroup of two or three languages which have at least one feature in common. The application of this methodology to the present dataset confirms Unruh & Kalisch' main intuitions concerning the internal genetic relationships within the Enlhet-Enenlhet language family. On the one hand, the subgroup that is most strongly supported by these data consists of Guaná, Enenlhet and Sanapaná, which has a subgroupiness value of 2,76 and excludes Enlhet, which, based on the limited Enxet data at my disposition, seems to be more similar to Enxet. This corresponds to Unruh & Kalisch' position that there is an east-west genetic divide within the family, the Eastern branch consisting of Angaité, Enenlhet, Guaná and Sanapaná, and the Western branch containing Enlhet and Enxet.

My data also confirm these authors' intuitions that the present-day Enlhet-Enenlhet languages stem from a dialect continuum rather than from a series of neat splits in a protolanguage. Rather than seeing a clean divide between Enlhet and the three other languages, Enlhet also shares a number of features with Guaná and Enenlhet, with Guaná and Sanapaná, and with Enenlhet exclusively. Within the Eastern Enlhet-Enenlhet subgroup, however, my data point in a different direction than those of Unruh & Kalisch. Rather than Guaná and Enenlhet grouping together, it seems that Guaná and Sanapaná are more similar with respect to the forms and functions of these three particles. From the few example sentences at my disposition, it seems that Enxet patterns more closely to Enlhet, and that Angaité seems to share features with all four languages. Nevertheless, I will not make any judgments about the position of these last two languages until more data become available.

7.2 Conclusions and suggestions for further research

In sum, this thesis has value for South American linguistics both from a descriptive point of view, and from a historical-comparative point of view. From a descriptive angle, it is the first study on a particularly fascinating aspect of an underdescribed language. It was concluded that the semantically very dense particles under scrutiny in this paper carry a TAM load and have predicative force, and that Sanapaná makes use of both a synthetic TAM system (with verbal realis, irrealis, and imperfective suffixes) and an analytic one (with the irrealis auxiliary which precedes the verb, and the particles that follow the verb).

Nevertheless, this is only the tip of the iceberg concerning Sanapaná (verbal) morphology. Many suffixal categories, expressing aspect (such as the complexive and the repetitive), direction and location (such as the ventive), and different modes of subordination (such as the subjunctive and the infinitive) remain to be described, as does the interaction of these categories with, for example, the different stems of verbs and the members of different verb classes. Furthermore, there are a number of other particles in Sanapaná, which phonologically and syntactically seem to have the same status as the one discussed here, but whose semantics belong to different domains (such as evidential marking). A study of these particles, and whether or not they behave in the same way (following nouns and verbs) and have any functional similarities (such as specifying the predicativity of nouns) would be a worthwile complement to the present investigation.

From a historical-comparative point of view, then, this work bundles observations on four Enlhet-Enenlhet languages. It provides us both with insights relevant to the structure of this language family, and it is a useful small-scale case study of the value of the wave model and historical glottometry for historical linguistics. The use of this methodology allows the tackling of problems that would be hard to get around within a traditional family tree model, such as the fact that even though Guaná, Enenlhet and Sanapaná group together most strongly, both Enenlhet and Guaná share features with Enlhet which they do not share with Sanapaná. Even though a basis was laid here for the reconstruction of the proto-Enlhet-Enenlhet dialect continuum, much more research, mainly of a descriptive kind, must be done on these languages (and especially on Angaité and Enxet) in order to successfully complete such an endeavour.

References

Alexiadou, Artemis (2008). Tense marking in the nominal domain: Implications for grammar architecture. *Linguistic Variation Yearbook* 8.1: 33-60.

Alvarsson, Jan-Åke (2007). The process of ethno[re]genesis among the 'Weenhayek of the Gran Chaco (1976-2006). *Revista del CESLA* 10: 139-56.

Barnes, Janet (1984). Evidentials in the Tuyuca verb. *International Journal of American Linguistics* 50.3: 255-71.

Bhat, D. N. S. (1999). *The Prominence of Tense, Aspect and Mood*. Amsterdam/Philadelphia: John Benjamins.

Beck, David (2013). Unidirectional flexibility and the noun-verb distinction in Lushotseed. In Jan Rijkhoff and Eva Van Lier (Eds.), 185-220.

Blake, Barry J. (1990). *Relational Grammar*. London: Routledge.

Booij, Geert (2006). Inflection and derivation. In Keith Brown (Ed.): *Encyclopedia of Language & Linguistics, Volume 5*, 654-61. Amsterdam: Elsevier.

Braunstein, José and Elmer S. Miller (1999). Ethnohistorical introduction. In Elmer S. Miller (Ed.): *Peoples of the Gran Chaco*, 1-22. Westport/London: Bergin & Garvey.

Brinton, Daniel G. (1898). The linguistic cartography of the Chaco region. *Proceedings of the American Philosophical Society* 37.158: 178-205.

Bybee, Joan (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.

Bybee, Joan L. and Östen Dahl (1989). The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13.1: 51-103.

Campbell, Lyle (2013). Language Contact and Linguistic Change in the Chaco. *Revista Brasileira de Linguística Antropológica* 5.2: 259-91.

Carrió, Cintia (2010). Los laberintos del tiempo y el hilo de Ariadna. La marcación de la temporalidad en la lengua mocoví (Familia guaycurú – Argentina). *Indiana* 27: 237-67.

Dahl, Östen (1985). *Tense and Aspect Systems*. New York: Blackwell.

de Haan, Ferdinand (2005). Typological approaches to modality. In William Frawley (Ed.): *The Expression of Modality*, 27-69. Berlin/New York: Mouton de Gruyter.

de Haan, Ferdinand (2011). Typology of tense, aspect and modality systems. In Jae Jung Song (Ed.): *The Oxford Handbook of Linguistic Typology*, 445-464. Oxford: Oxford University Press.

Deo, Ashwini (2012). Morphology. In Robert I. Binnick (Ed.): *The Oxford Handbook of Tense and Aspect*, 155-183. Oxford: Oxford University Press.

DGEEC: Dirección General de Estadística, Encuestas y Censos (2012). III Censo Nacional de Población y Viviendas para Pueblos Indígenas 2012. Resultados Finales. Asunción.

Durante, Santiago (2011). Las lenguas del Gran Chaco: Situación socio-lingüística y políticas lingüísticas. *Language design: journal of theoretical and experimental linguistics* 13: 115-42.

Evans, Nicholas (2000). Word classes in the world's languages. In Geert Booij, Christian Lehmann and Joachim Mugdan (Eds.): *Morphology: A Handbook on Inflection and Word Formation*, 708–732. Berlin: Mouton de Gruyter

Evans, Nicholas and Toshiki Osada (2005). Mundari: The myth of a language without word classes. *Linguistic Typology* 9.3: 351-90.

Foley, William (1991). *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.

François, Alexandre (2014). Trees, waves and linkages: Models of language diversification. In Claire Bowern and Bethwyn Evans (Eds.): *The Routledge Handbook of Historical Linguistics*, 161-89. London/New York: Routledge.

Frawley, William (1992). *Linguistic semantics*. Hillsdale, NJ: Lawrence Erlbaum.

Givón, T. (2001). *Syntax, Volume 1*. Amsterdam/Philadelphia: John Benjamins.

Golluscio, Lucía and Alejandra Vidal (2009). Recorrido sobre las lenguas del Chaco y los aportes a la investigación lingüística. *Amerindia* 33/34: 3-40.

Gomes, Antônio Almir Silva (2009a). Sanapaná: Um povo indígena paraguaio. *Sínteses* 14: 360-85.

Gomes, Antônio Almir Silva (2009b). Notas sobre a (morfo)fonologia Sanapaná. Paper presented at the *Workshop da Área de Línguas Indígenas*. Campinas, Universidade Estadual de Campinas, 12-13 November 2009.

Gomes, Antônio Almir Silva (2011). Gênero e número em Sanapaná. *Anais do VII Congreso Internacional da Abralin*, 528-39.

Gomes, Antonio Almir Silva (2013). Sanapaná uma língua Maskoy: Aspectos gramaticais. Doctoral dissertation, Universidade Estadual de Campinas.

Gregores, Emma and Jorge Suárez (1967). *A description of colloquial Guaraní*. The Hague: Mouton.

Grimes, Barbara (2003). *Ethnologue: Languages of the World (online version)*. Fourteenth version, January 2003. SIL International.

Hammarström, Harald, Robert Forkel and Martin Haspelmath (2017). *Glottolog 3.0*. Jena: Max Planck Institute for the Science of Human History. Accessed 16 May 2017. http://glottolog.org.

Haude, Katharina (2009). Reference and predication in Movima. In Patience Epps and Alexandre Arkhipov (Eds.): *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, 323-42. Berlin: Mouton de Gruyter.

Hengeveld, Kees (1992). Parts of speech. In Michael Fortescue, Peter Harder and Lars Kristofferson (Eds.), *Layered Structure and Reference in a Functional Perspective*, 29-53. Amsterdam/Philadelphia: John Benjamins.

Hengeveld, Kees, Jan Rijkhoff and Anna Siewierska (2004). Parts-of-speech systems and word order. *Journal of Linguistics* 40: 527-70.

Hopper, Paul J. (1991). On some principles of grammaticization. In Elizabeth C. Traugott and Bernd Heine (Eds.): *Approaches to Grammaticalization, Volume 1*, 17-36. Amsterdam/Philadelphia: John Benjamins.

Jelinek, Eloise and Richard A. Demers (1994). Predicates and pronominal arguments in Straits Salish. *Language* 70.4: 697–736.

Kalisch, Hannes (2009). Los constituyentes de la cláusula enlhet (enlhet-enenlhet). Esbozo de una cláusula omnipredicativa. *Amerindia* 33/34: 109-50.

Kalisch, Hannes (2012). Lecturas de las fiestas Enlhet. In Ticio Escobar (Ed.): *La Belleza de los Otros*, 349-65. Asunción: Servilibro.

Kidd, Stephen W. (2000). Love and hate among the people without things: The social and economic relations of the Enxet people of Paraguay. Doctoral dissertation, University of St. Andrews.

Koshal, Sanyukta (1979). *Ladakhi Grammar*. Delhi: Motilal Banarsidass.

Krebs, Edgardo and José Braunstein (2011). The renewal of Gran Chaco studies. *History of Anthropology Newsletter* 38.1: 9-19.

Launey, Michel (1994). *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique*. Paris: CNRS.

Launey, Michel (2004). The features of omnipredicativity in Classical Nahuatl. *Sprachtypologie und Universalienforschung* 57.1: 49-69.

Lecarme, Jacqueline (1999). Nominal tense and tense theory. In Francis Corblin, Carmen Dobrivie-Sorin and Jean-Marie Marandin (Eds.): *Empirical Issues in Formal Syntax and Semantics 2*, 333-54. The Hague: Thesus.

Lemaréchal, Alain (1989). *Les parties du discours. Sémantique et syntaxe*. Paris: Presses Universitaires de France.

Macaulay, Monica Ann (1996). *A Grammar of Chalcatongo Mixtec*. Berkeley: University of Los Angeles Press.

Majewicz, Alfred F. (1985). *The Grammatical Category of Aspect in Japanese and Polish in a Comparative Perspective*. Poznan: Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza.

Merriam-Webster (2017). Predicate. Web: Merriam-Webster.com. Consulted 16 May 2017.

Nordlinger, Rachel and Louisa Sadler (2004a). Nominal tense in crosslinguistic perspective. *Language* 80.4: 776-806.

Nordlinger, Rachel and Louisa Sadler (2004b). Tense beyond the verb: Encoding clausal tense/aspect/mood on nominal dependents. *Natural Language & Linguistic Theory* 22.3: 597-641.

Palmer, Frank R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.

Parker, Steve (1999). On the behavior of definite articles in Chamicuro. *Language* 75.5: 52-62.

Perlmutter, David M. (1980). Relational grammer. In Andrew Moravcsik and J.R. Wirth (Eds.): *Syntax and Semantics: Current Approaches to Syntax, Vol. 13*, 195-229. New York: Academic Press.

Pinker, Steven and Paul Bloom (1990). Natural language and natural selection. *Brain and Behavioral Sciences* 13: 707-84.

Queixalós, Francesc (2006). The primacy and fate of predicativity in Tupi-Guarani. In Ximena Lois and Valentina Vapnarsky (Eds.): *Lexical Categories and Root Classes in Amerindian Languages*, 249-88. Bern: Peter Lang.

Rathert, Monika and Renate Musan (2012). *Tense across Languages*. Berlin: Mouton de Gruyter.

Reichenbach, Hans (1947). Elements of symbolic logic. New York: the Free Press; London: Collier/Macmillan.

Rijkhoff, Jan and Eva Van Lier (2013). *Flexible Word Classes: Typological Studies of Underspecified Parts of Speech.* Oxford: Oxford University Press.

Sasse, Hans-Jürgen (1991). Aspect and aktionsart: A reconciliation. *Belgian Journal of Linguistics* 6.1: 31-45.

Sasse, Hans-Jürgen (2006). Aspect and aktionsart. In Keith Brown (Ed.): *Encyclopedia of Language and Linguistics*, 535-38. Amsterdam: Elsevier.

Schmidt, Johannes (1872). *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen.* Weimar: Hermann Böhlau.

Siewierska, Anna (1991). *Functional Grammar*. London: Routledge.

Siewierska, Anna (1998). On nominal and verbal person marking. *Linguistic Typology* 2.1: 1-56).

Stowell, Tim (2012). Syntax. In Robert I. Binnick (Ed.): *The Oxford Handbook of Tense and Aspect*, 184-211. Oxford: Oxford University Press.

Sušnik, Branislava (1962). *Estudios Emok-Toba*. Asunción: Museo Etnográfico Andrés Barbero.

Sušnik, Branislava (1977). *Lengua-Maskoy. Su Hablar – Su Pensar – Su Vivencia*. Asunción: Museo Etnográfico Andrés Barbero.

Swadesh, Morris (1939). Nootka internal syntax. *International Journal of American Linguistics* 9: 77-102).

Tonhauser, Judith (2007). Nominal tense? The meaning of Guaraní nominal temporal markers. *Language* 83.4: 831-69.

Tonhauser, Judith (2008). Defining cross-linguistic categories: The case of nominal tense (reply to Nordlinger and Sadler). *Language* 84.2: 332-42.

Unruh, Ernesto and Hannes Kalisch (1999). *Vana Peema: Nentenyey'a Nelhpaqmeesma Nempeema; guía del Maestro Maskoy para el Aprendizaje del Idioma Guaná.* Ya'alve Sanga: Nengvaanemquescama Nempayvaam Enlhet.

Unruh, Ernesto and Hannes Kalisch (2003). Enlhet-Enenlhet. Una familia lingüística chaqueña. *Thule, Rivista italiana di studi americanistici* 14/15: 207-31.

Unruh, Ernesto; Hannes Kalisch and Manolo Romero (2003). *Enenlhet apaivoma. Nentengiai'a nengiangveikmoho neliatekama enenlhet apaivoma. Guía para el Aprendizaje del Idioma Enenlhet-Toba.* Ya'alve-Saanga: Nengvaanemkeskama Nempayvaam Enlhet.

van der Auwera, Johan and Andreas Ammann (2013). Overlap between situational and epistemic modal marking. In Dryer, Matthew S. and Martin Haspelmath, Matthew Dryer, (Eds.): *World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed 20 June 2017.

Villagra Carrón, Rodrigo (2010). The Two Shamans and the Owner of the Cattle: Alterity, Storytelling, and Shamanism amongst the Angaité of the Paraguayan Chaco. Doctoral dissertation, University of St. Andrews.

Zanardini, José (2012). *Lenguas indígenas del Paraguay. In Ministerio de Educación, Cultura y Deporte and Secretaría General Técnica (Eds.): Actas del Congreso Iberoamericano de las Lenguas en la Educación. Las Lenguas en la Educación, Cine, Literatura, Redes, y Nuevas Tecnologías*, 35-42. Madrid: Ministerio de Educación, Cultura y Deporte.