8-1-2015

Case and pragmatic status markers in Embera Katío

Bethany Winter Carlson

CASE AND PRAGMATIC STATUS MARKERS IN EMBERA KATÍO

by

Bethany Winter Carlson
Bachelor of Arts, Eastern Michigan University, 2012

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Arts

Grand Forks, North Dakota
August
2015

This thesis, submitted by Bethany Winter Carlson in partial fulfillment of the requirements for the Degree of Master of Arts from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

_____

Dr. John M. Clifton

_____

Dr. Stephen A. Marlett

_____

Dr. Keith Slater

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

_____

Dr. Wayne E. Swisher
Dean of the School of Graduate Studies

_____

Date

PERMISSION

Title	Case and Pragmatic Status Markers in Embera Katío
Department	Linguistics
Degree	Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Bethany W. Carlson

July 2, 2015

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

1	first person
2	second person
3	third person
ABL	ablative
AUX	auxiliary
BEN	benefactive
CAUS	causative
CLF	classifier
COMPL	completive
COND	conditional
DAT	dative
DECL	declarative
DEF	definite
DIM	diminutive
EMPH	emphasis
ERG	ergative
EVID	evidential
EXCL	exclusive
FOC	focus
FUT	future
HAB	habitual
IMP	imperative
INCP	inceptive
INCL	inclusive

INS	instrumental
IRR	irrealis
LOC	locative
NEG	negation, negative
NONVOL	nonvolitional
NPRS	nonpresent
OBJ	object
OBLG	obligation
ORIG	origin
PL	plural
POSS	possessive
PRS	present
PROG	progressive
PST	past
PURP	purposive
Q	question particle/marker
REAS	reason
REPT	repetitive
SBJ	subject
SEQ	sequential
SG	singular
STV	stative
TOP	topic

ACKNOWLEDGMENTS

My work on this thesis was only possible because of much support, input and advice from others.

First, I want to offer special thanks to those Embera brothers who attended our men's workshops and willingly shared their language and stories with me, patiently answering my many strange questions and asking their own questions about my language, family and country. Thank you to Luis Domicó Domicó, Mario Majoré, Jose Aurelio Sapia, Adelmo Ciro Sapia Casamá, José Domicó Domicó, Miguel Antonio Domicó Domicó, Jose Eligio Majoré Casamá, Jhon Fredi Domicó Domicó and Jhonneiro Domicó Casama. Very special thanks to my language consultant Josué, whose name has been changed because he does not wish to be named, but who was most vital to my research and was kindly available to answer my questions anytime.

I give a special thank you also to Gisella Greenfield, my mentor and dear friend, for sharing her life and home with me in these last three years. I am indebted to her many years of dedicated work with the Katíos that opened the doors of trust for me to enter their community. Her advice and help have been important to my research, and her support and love have been invaluable to me personally. Thank you to Eliana Kerguelén, whose love for the Katíos is contagious, for her help with my research on several occasions and for providing me with a copy of the Katío dictionary (Villalobos & Garcia 2004) that became a major source of texts for this thesis.

A very special thank you to John Clifton, Steve Marlett and Keith Slater for their very helpful comments and guidance in the writing and editing process and for taking the time to make my thesis better and challenge my thinking.

Thank you to Diana Weber for her help proofreading and her insightful suggestions, and to David Weber and Linda Humnick for their valuable input.

I would like to express my gratitude to Chaz Mortensen for his helpful explanation about the orthography and his willingness to elaborate on his research by personal correspondence, and to Phill Harms for his correspondence. I'm grateful also to David and Penelope Pickens for opening their home to us and generously sharing linguistic and literacy materials with us, and to them and their Embera coworkers Frank, Liam and Carlos for explaining some aspects of Embera Baudó and Embera Tadó.

Finally, thank you to my dear family, who has always supported me unconditionally.

ABSTRACT

In previous research on the Embera languages of the Chocó genus of Colombia and Panama, many different terms have been applied to one or all of the morphemes *ba*, *ta* and *ra* that frequently mark noun phrases and dependent clauses, and some of these descriptions have been conflicting. This study summarizes the previous analyses and compares them to language data from a corpus of texts and elicited material. The result is an analysis in which *ba* has three functions, marking ergative case, instrumental case and reason; *ta* marks focus on absolutives; and *ra* marks pragmatic topic and syntactic topics. In addition to conflicting descriptions of their functions, some researchers call them suffixes and some call them enclitics. This disagreement stems from the fact that, like prototypical affixes, they are part of the prosodic word, subject to nasal spreading, affect the stress of the preceding word and even bear stress. However, like prototypical clitics, they mark phrases and clauses, not words.

CHAPTER 1

INTRODUCTION

The Embera Katío are a peaceful semi-nomadic people who live in the mountains and along the rivers of northwest Colombia. They build their homes isolated from each other in nuclear family groups, and live by fishing, farming, and working for local ranchers. Since many live in remote areas, their language is still strong and there are many monolingual speakers.

I have worked in Colombia since 2012, studying the Katío language and assisting a team that has been working with the Katío people for over a decade. While the political situation of the region of the country where the Katíos live has not allowed me much access to their communities, I have learned about their language and culture from the Katío men who attend our bi-annual events, and from the community of Katío families in El Pindo in Montelíbano, Córdoba.

This study describes the functions of the morphemes *ba*, *ta* and *ra* in Embera Katío. They have been described in previous research with widely varying terminology, and this study summarizes those previous analyses, compares them to the function and distribution observed in a corpus of texts, and moves toward a unified term and analysis for each morpheme that describes its function.

Chapter 1 places Embera Katío in its linguistic family and geographic setting in 1.1, introduces what previous research has said about *ba*, *ta* and *ra* in 1.2, presents the arguments for an unmarked absolutive case in 1.3, discusses the texts and methodology used for this study in 1.4, and describes the orthography, the process of nasal spreading, and gives a brief introduction to the syntax of the language in 1.5.

Chapter 2 describes the ergative case marker *ba* and its other functions. Chapter 3 discusses *ta* and the concept of focus, the distribution of *ta* as a focus marker of absolutives

and complement clauses, and other morphemes of marked focus. Chapter 4 discusses *ra* and the concept of topic, and the function of *ra* in marking the pragmatic status of topicality and syntactic topics of topic-comment structures.

Chapter 5 presents the affix-like and clitic-like features of these morphemes and discusses the issues with calling them either suffixes or clitics.

Finally, chapter 6 summarizes the results of the study, briefly mentions some implications for the typological classification of the Embera languages in the World Atlas of Language Structures Online (Dryer & Haspelmath 2013), and suggests areas for further research.

1.1 Embera and the Chocó family

The Embera languages belong to the Chocó genus and are spoken mainly in Colombia and to a lesser extent in Panama (Lewis et al 2014). Chocó is currently considered an isolate language family, although it has previously been classified by different researchers as part of various larger families. Greenberg classed the family as Chibchan-Paezan (Huber & Reed 1992:xi), and it has also previously been considered to be Carib or Chibchan (Loewen 1963b:244). Costenla and Margery present some possible evidence for a relation with Chibchan (1991:172).

However, none of these classifications has yet been proven conclusive. Greenfield (2012:4) cites Aguirre Licht (2009), saying that the isolate classification is currently viewed as the "most reliable." She also states that the Colombian government recognizes the language family as independent. Pardo and Aguirre Licht (1993:292), who undertake a close examination and rebuttal of Rivet's (1943) classification of Chocó as a relation of Carib, are also in agreement with the classification of the Chocó group as an "independent linguistic family."

The Chocó family is made up of two groups: Waunana (also called Wounaan or Wounmei) and the Embera branch (Campbell 1997:172 cites Gunn 1980). Waunana is a single language with three regional dialects (Mejía Fonnegra 2000:57), while Embera is a subfamily, or "dialect continuum" (Adelaar 2004:57) which is divided into two main groups:

Northern Embera and Southern Embera. There is almost no mutual intelligibility between these two groups (Loewen 1963b:243).

Northern Embera includes two languages: Northern or Darién Embera (also called Embera and Northern Embera Proper) and Embera Katío, the focus of this thesis. In this study I follow Mortensen's (1999) convention of referring to the Northern Embera language as Northern Embera Proper (NEP) in order to distinguish it from the Northern Embera group.

Southern Embera includes four languages: Embera Chamí, Epena Pedee (also called Epena Saija), Embera Baudó and Embera Tadó (Mortensen 1999:1).

It is possible that the number of languages in the Embera family may be greater than the six mentioned above, however, and there is need for survey work and historical analysis to back up these groupings. Pardo and Aguirre Licht (1993:298) suggest seven mutually unintelligible varieties of Embera arranged in five groups which they define based on the phonemic inventory (1993:297). In 1963, Loewen identified at least nine dialectal areas of Embera (1963b:241) based on "distinguishing features from the phonology, morphology and the lexical inventory of the respective dialects" (1963b:243).

Embera Katío (ISO 639-3:[cto]) is also called by the names Catío, Eyabida, and Dabeiba. It is spoken mainly in northwestern Colombia in the departments of Córdoba, Chocó, and Antioquia along the upper Sinú, San Jorge, San Pedro, and Murrí rivers (Lewis et al 2014). According to the most recent Colombian national census in 2005, the Embera Katío population numbers 38,259, with 33.5% in Antioquia, 26.5% in Chocó, 13.4% in Córdoba, and 13.6% in urban areas.

Figure 1 shows the approximate geographic territory occupied by the Embera Katío people group, as well as two neighboring groups, the Northern Embera Proper (NEP) and the Kuna.

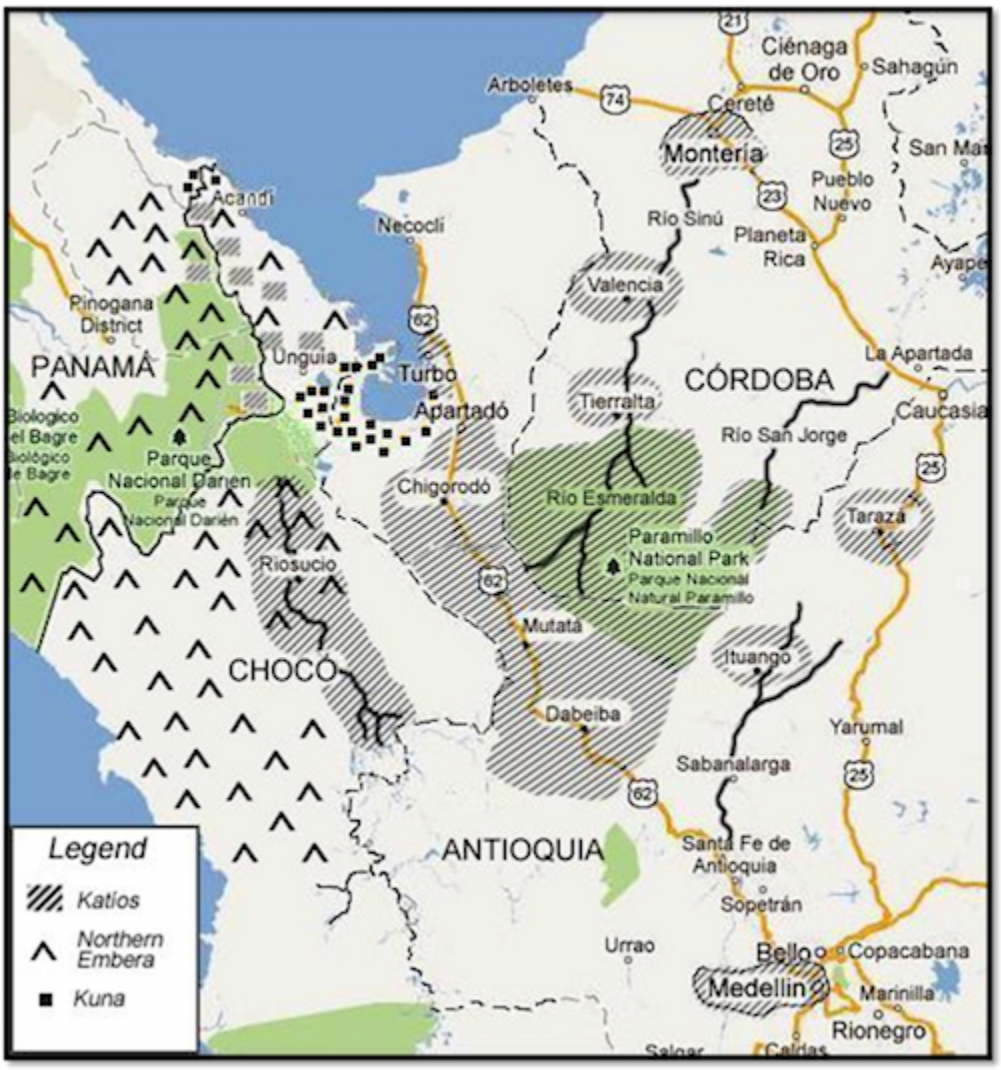

Figure 1. The territory of the Embera Katío people[1]

However, not all who identify themselves as Katíos are speakers of the Embera Katío language. My colleague and I met a group of Embera families who had arrived in the city of Cali after being displaced from their lands in the Colombian department of Risaralda in the territory where the Southern Embera languages are spoken. The women's dress was in the style of the Southern Embera. They identified themselves as Katíos, yet their speech was so different from the Embera Katío language of Córdoba, Chocó and Antioquia that communication in Embera was not possible.

[1] The map in figure 1 is from Greenfield (2012:2), used with permission.

This was confirmed by some Embera Tadó acquaintances, who in a personal conversation reported that they knew of Embera of other languages calling themselves by the name Katío. It may be that the number of self-reporting Katíos is greater than the number of speakers of this particular language. However, when I use the term in this study, it refers only to the speakers of the Embera Katío language.

The name Katío is an exonym; the Embera Katíos themselves call their language *Embera Bedea* or 'language of the people'. Since many Embera peoples call themselves a variation of this, I will use the term Embera Katío or Katío in order to distinguish between them. The Embera Katíos use the word Embera, written *Ẽbẽra* in the orthography, to refer to themselves and to other indigenous peoples. Although it literally translates as 'person', *ẽbẽra* is used to distinguish between indigenous peoples and majority Latino Colombians, whom they call *k'apũrĩã*. They call the Spanish language *K'apũrĩã Bedea* or 'language of the non-indigenous people'.

1.2 Previous research

The Embera Katío language has a group of morphemes that mark grammatical relations and pragmatic status.[2] I examine three of these in this study: *ba*, *ta* and *ra*. Some researchers refer to these as suffixes, and some as enclitics. In reality they have features associated with both, which I discuss in chapter 5.

I cite six researchers that have mentioned these morphemes in their publications on Embera languages. The first is Jacob Loewen (1958), who described the Sambú dialect of Northern Embera in southern Panama in his dissertation. Mareike Schöttelndreyer, who lived and worked with the Katío people for years, mentions these morphemes in passing in her analysis of a Katío folkloric narrative (Schöttelndreyer & Levinsohn 1976). Her coworker, Eileen Rex, wrote her thesis on the grammar of Embera Katío for the University of Texas at Arlington (1975). Phillip Harms wrote a grammar of Epena Pedee, a Southern Embera language spoken along the southern part of the Pacific coast of Colombia (1994). Daniel Aguirre Licht (1999), a Colombian anthropologist and ethnolinguist of Andes University in Bogotá and author of numerous books on the Embera language family, wrote a

[2] The term pragmatic status comes from Payne's chapter on pragmatically marked structures (1997:261).

short article on Embera Chamí, another Southern Embera language. Charles Mortensen (1999) models his reference grammar of the Northern Embera languages on Harms (1994).

The three morphemes of this study have been described differently by the researchers above, reflecting disagreement as to their category and function. The following is a brief overview of the different analyses. I discuss the previous research on each morpheme in more detail at the beginning of its corresponding chapter.

There is general consensus that *ba* marks ergative case, but the terms "agentive" and "ablative" have also been applied to *ba*, and are discussed in section 2.2. Rex describes an enclitic *ba* that marks ergative case, instrument and the "locative of source" (1975:38). Mortensen also describes a suffix *ba* that marks ergative and instrument, but instead of talking about source, he adds that it marks reason (1999:47). Chapter 2 discusses the various functions of *ba*.

The descriptions of *ta* and *ra*, on the other hand, are more divergent than those of *ba*. They are both called "enclitic emphatic particles" by Loewen (1958:101). Rex considers *ta* to be an enclitic that marks the equivalent of absolutive case (1975:39), and Harms describes *ta* in Epena Pedee as a focus enclitic (1994:193). Mortensen combines both ideas and calls it an "absolutive focus suffix" (1999:143).

As for *ra*, Rex calls it a "non-case marker" of "previous reference" (1975:39-40), while Aguirre Licht describes a topic marker suffix *ra* in Embera Chamí (1999:327), and Mortensen (1999) considers it to be the "normal" or "nonfocal" absolutive marker suffix (1999:143, 49).

With *ta* and *ra*, the disagreement in the analyses includes function (are they case markers, and do they mark emphasis, previous mention, focus, or topic?) and whether they are suffixes or enclitics. Their functions are discussed in chapters 3 and 4 respectively, and the difficulty with calling them simply suffixes or enclitics is presented in chapter 5. In the next section I will briefly discuss the claims that they are absolutive case markers.

1.3 Absolutive case

In this section, I lay the groundwork for the discussion of *ta* and *ra* by arguing that absolutive case is unmarked and *ta* and *ra* are not absolutive case markers.

Table 1 summarizes the analyses of five of the six researchers[3] presented in the previous section as to the marking of the absolutive case and the functions of *ta* and *ra*.

Table 1. Summary of the analyses of *ta* and *ra*

	Absolutive Case	**Pragmatic Status Marker**
Loewen	unmarked	*ta / ra*
Rex	*ta*	*ra*
Harms	unmarked	*ta*
Aguirre Licht	unmarked	*ra*
Mortensen	*ra*	*ta*

There is agreement that the absolutive case is unmarked in the Southern Embera languages Embera Chamí (Aguirre Licht 1999:317-318) and Epena Pedee (Harms 1994:66). Although Loewen does not write in terms of case, his description stating that *ta* and *ra* mark emphasis implies an analysis of unmarked absolutive case since there are no other morphemes that appear on absolutives. I suggest that this is also a reasonable analysis for Embera Katío's absolutive case.

As noted above, Rex describes *ta* as an absolutive case marker (1975:39), and Mortensen classes *ra* as "normal absolutive case marking" (1999:143). However, calling them case markers fails to account for the high frequency of absolutives that are unmarked, as in examples (1), (3) and (4b).

In example (1), the subject *mɨ* '1SG' of the intransitive verb *wã* 'go' is unmarked, even though *ra* can also occur with the pronoun *mɨ*, as shown by example (2).[4]

(1) *Ewari aba wã-si-a* ***mɨ*** *Sarroma* =*ne ena.*
day one go-PST-DECL 1SG Sarroma =LOC LOC
One day I went to Sarroma. [Hammock:1][5]

(2) *Mɨ* =*ra nuweda ze-si-a.*
1SG =TOP yesterday come-PST-DECL
I came yesterday.

[3] The mentions found in Schöttelndreyer (1977) are briefly referencing the interaction of *ta* and *ra* with the conflict in the narrative, and only a cursory linguistic description of them is given in a final footnote, which is a reference to Rex (1975).

[4] In the examples, I present morphemes that are not affixes separated with a space from their host words with an = sign before to illustrate phonological dependence.

[5] The key for cited texts is shown in table 2 in section 1.4. Examples with no source given are from elicited material.

The object noun phrase *opoga zhara* 'iguana meat' is unmarked in example (3).

(3) *Ẽbẽra =ba **[opoga zhara]** k'o-ba-d'a, kũruma wia k'o-ba-d'a, pada*
person =ERG iguana meat eat-HAB-PL smoke cook eat-HAB-PL plantain

ed'a k'o-ba-d'a.
LOC eat-HAB-PL
Embera eat iguana meat, they eat it smoked and cooked, they eat it with plantain.
[V&G:p157]

Example (4a) establishes *mɯ* '1SG' as the ergative argument of the verb. In (4b), *k'abayo* 'horse' (borrowed from Spanish), the object of the transitive verb *zrɯga* 'steal', is unmarked.

(4) a. *Me mɯ =a zrɯga-ya.*
alright 1SG =ERG steal-FUT
Alright, I'll steal [it].

b. ***K'abayo** zrɯga-ya.*
horse steal-FUT
I'll steal the horse. [EP: 106, 107]

Mortensen accounts for the unmarked object by saying that in Embera Katío, "direct objects don't require *ra*" (personal communication). This could account for examples (3) and (4b), but fails to account for example (1) in which *mɯ* '1SG' cannot be a direct object since it is not a patient of the verb *wã* 'go'.

Mortensen also states in his grammar that an absolutive case noun receives zero marking when it is nonactivated, that is, when it has "not yet been activated or brought into focus in a text," which could account for example (1), but not for (4b) in which *k'abayo* is already activated in the discourse (see example (100)). In fact, Mortensen posits four absolutive suffixes, depending on the activation status of the absolutive: *-ø*, *-ra*, *-ta*, and *-trɯ*, which are nonactivated, "normal" activated and nonfocal, introductory focus, and focus on given information respectively (1999:143). According to Mortensen's analysis, case and the activation status coding are fused into a single morpheme.

My analysis separates the activation status or what I more generally call pragmatic status from the absolutive case. I consider absolutive case to be unmarked, and *ta* and *ra* to be independent of case, carrying only their respective pragmatic information.

A significant problem with the analysis of *ra* as an absolutive case marker is the frequent occurrence of *ra* together with ergative *ba* as in example (5) and other case markers like the benefactive as in (6), as well as with other syntactic categories which cannot be absolutives.

(5) *Bok'orro* =***ba*** =***ra*** *k'o-ba-ri-a* *egoro* =*ta...*
frog =ERG =TOP eat-NPRS-HAB.PRS-DECL earth =FOC
The frog eats dirt. [V&G:p42]

(6) *Do* =*ra* *mɯ* =***ita*** =***ra*** *mɯ* *mebea* *kĩrãk'a* *b'-ɯ-a.*
river =TOP 1SG =BEN =TOP 1SG sibling like AUX-PRS-DECL
The river, for me [it] is like a sibling.

Mortensen accounts for this by positing two morphemes with the form *ra*, but I argue that they are just one morpheme in section 4.3.

I propose that Mortensen's assertion that *ta* marks "focus" and Rex's assertion that *ra* marks "previous reference" are steps in the right direction, and in this study I flesh out those functions by elaborating on the notion of "focus" and using the term topic to encompass "previous reference." My analysis is similar to Loewen's in table 1, except that Loewen does not describe the functions of *ta* and *ra* in detail nor does he distinguish between them.

Considering the arguments above and the fact that typologically, in ergative systems, absolutive case is most commonly unmarked (Givón 1984:151), I conclude that unmarked absolutive case is the best analysis for Embera Katío as well. Thus *ta* and *ra* are not absolutive case markers, but rather *ta* marks focus, discussed in section 3, and *ra* marks topic, discussed in section 4.

1.4 Data corpus and methodology

My methodology for this study included elicitation sessions with native-speaker language consultants from Dabeiba, Antioquia and from Alto San Jorge, Córdoba, as well as analysis of a corpus of texts and elicited sentences.

The data in the examples in this thesis are from a body of texts collected from published works and my own research. Also included are lists of 380 isolated elicited sentences, some collected by my colleague Gisella Greenfield, and some by myself. All examples without a cited source text are from elicited material.

Three collections of example sentences were elicited, transcribed and translated by Greenfield with various language consultants, and two collections of example sentences were elicited and transcribed by myself with the help of Josué (not his real name), a speaker from Antioquia who wishes not to be identified. In these sentences, some aspects such as the definiteness of the referents may be ambiguous in certain cases, and slightly varying free translations might be possible.

Table 2 is a summary of the texts referenced in this study.

Table 2. Texts and their abbreviations and sources

Text title	**Abbreviation**	**Source**
Africanas	-	Personal fieldwork
Gossip	-	Personal fieldwork
Fight	-	Personal fieldwork
Hammock	-	Personal fieldwork
Prayer	-	Personal fieldwork
23 short texts	V&G	Villalobos & Garcia (2004)
Feliz Duerto	FD	Schöttelndreyer (1977)
El Pegante	EP	Schöttelndreyer (1977)
Long personal narrative text	BBT	Mortensen (1999)

The texts in table 2 are cited with sentence numbers except for the texts in Villalobos & Garcia (2004), which are cited with page numbers.

The texts *Africanas*, *Gossip*, *The Fight*, *Hammock*, and *Prayer* come from personal fieldwork, which means they were elicited, transcribed and translated by myself and my colleagues Gisella Greenfield and Eliana Kerguelén with the help of our language consultants Luis Domicó Domicó, Mario Majoré, Jose Aurelio Sapia, Adelmo Ciro Sapia Casamá, José Domicó Domicó, Miguel Antonio Domicó Domicó, Jose Eligio Majoré Casamá, Jhon Fredi Domicó Domicó and Jhonneiro Domicó Casamá. These texts were elicited during several of the bi-annual men's workshops hosted by our team in 2013 and transcribed and translated over the course of several consecutive workshops. Each participant was asked to

share a story with the group about something that happened to himself or someone he knew. Altogether, we collected recordings that totaled 45 minutes, 26 minutes of which we transcribed and translated.

The texts *Ãparra*, *Bakuru*, *Beda sãwa bea bada?*, *Bẽrogõara sãwa bea bada?*, *Bido*, *Bokorro*, *Bõpa*, *Dazhi be bari de eda bema*, *De sãwa o bada?*, *Dora*, *Drũdaʉ*, *Ẽbẽra de*, *Euma nebʉra*, *Jedeko*, *Jenenera*, *Kẽberre*, *Kõe Kõe*, *Kore*, *Korogo*, *Kuriwa*, *Nusi*, *Opoga* and *Warru* all come from the 2004 Ethnolinguistic Dictionary of the Embera of Alto Sinú published by Villalobos and Garcia (2004). These original texts were published as a cultural record of the Embera Katíos of the upper Sinú river, but included no translation, interlinearization or analysis, so all translation and analysis is my own with the help of my language consultant Josué. I adjusted the orthography where necessary for consistency. The length of these texts varies from 3 to 63 sentences, totaling 357 sentences. The texts listed above include only those cited in this thesis as examples, however, as part of the study I interlinearized and at least partially translated all 84 texts published in the dictionary.

The texts *Feliz Duerto* (46 sentences), narrated by Yadibi Bailarín, and *El Pegante* (161 sentences), narrated by Angel Bailarín, were published as appendixes in Schöttelndreyer (1977). The texts published by Schöttelndreyer use an outdated orthography, so in my citation I have updated the orthography for consistency (using *k* instead of Spanish-influenced *c* and *qu*, for example). Additionally, these texts were published with word glosses only and have no analysis of individual morphemes, so the interlinearization is in great part my own. In some cases I have also adjusted the free translation for clarity and accuracy.

I cite a text of 140 sentences that is published as an appendix in Mortensen (1999) and narrated by Abel Domicó, and I also cite selected example sentences from the same work. In the case of Mortensen's reference grammar, where the text was published in interlinearized form with a free translation, my analysis differs from Mortensen's in some points regarding the nature and glossing of certain morphemes, particularly *ra*. In citing examples from the text published in Mortensen, I have glossed according to my own analysis, referring to Mortensen's original glossing where relevant. For this text too, I adjusted the

orthography for consistency since the text is published with a phonemic representation rather than an orthographic representation as I have used.

The purpose of this study is to accurately describe the case marking and pragmatic status systems of Embera Katío in order to document these features and contribute to the data which form the basis of typological generalizations. Thus I am less interested in applying or testing a formal theoretical framework against the data, and more interested in creating "a description that is consistent with the properties of the language itself" (Payne 1997:3). I hope that my description will allow Embera Katío and the Embera family to be more accurately mapped typologically. My framework is functional typological following Givón (1984) and Payne (1997).

I agree with Haspelmath (2010) that working from a basis of crosslinguistic or formal categories which we then impose upon the structures of individual languages is unhelpful in a world where each language is structured uniquely with its own uniquely-defined categories. Of course, in order to meaningfully compare languages, we must have a set of "comparative concepts" on which to draw, but these should not be confused with the language-specific categories which will not line up exactly with the scope of the comparative concept. I therefore borrow terminology associated with crosslinguistic comparative concepts in order to name language-specific descriptive categories without presuming that all that is encompassed in the prototypical dative case, to choose an example, will be true for the dative case of Embera Katío or vice versa.

1.5 Basic features of Katío

1.5.1 Orthography

In this study I present all syntactic examples in orthographic notation. As noted above, I have made orthographic adjustments in the language data for consistency.

Katío phonology has been an area of disagreement among researchers, particularly in the definitions of the three series of stops (Rex & Schöttelndreyer (1972) posit two.) This is perhaps due to the "continuum" nature of the Embera language family, the variation between dialects and speakers regarding the pronunciation of the series of stops, as well

as the fact that the phonetic realization of the stops can be affected by the process of nasal spreading described in 1.5.2. Thus the phonetic realization of the stop series, especially the voiced series, may vary in different dialects with regards to voicing.

The orthography of Embera Katío reflects three stops for the bilabial and alveolar places of articulation and two stops for the velar place of articulation. The stops *p*, *t* and *k* are slightly aspirated, the stops *b*, *d* and *g* are voiced, and the stops with apostrophes *b'* and *d'* represent implosives [ɓ] and [ɗ] respectively.[6]

The orthography marks an allophonic process in which a velar stop is backed and either affricated or spirantized when it is adjacent only to back vowels *a* and *o* and not to high vowels *i*, *e*, *u* or *ɯ*. Thus, the two velar stops *k* and *g* appear marked with an apostrophe *k'* and *g'* when they predictably undergo this process. The spirantization is stronger in the Katío of Dabeiba and less marked in the Katío of Alto San Jorge.

Diagraphs *ch* and *ʑh* represent affricates [tʃ] and [dʒ] respectively, *y* represents palatal approximant [j], *j* represents [h], *rr* alveolar trill [r], and *r* alveolar flap [ɾ]. All other consonants in the orthography represent their IPA phonetic values.

The letters *ʉ* and *ʉ̃* represent high back unrounded vowels [ɯ] and [ɯ̃] respectively. The other vowel letters are roughly equivalent to their phonetic values.

1.5.2 Nasal Spreading

One of the most notable features of Katío phonology is its process of nasal spreading, or the spreading of the nasal feature from a nasal element to surrounding segments. This is reflected in the variety of spellings of the Embera language family name by different researchers and in different Embera languages, trying to account for the nasalization of the vowels and of *r* and the prenasalization of *b*: Embera, Ẽberã, Epena, Embena, Ebera, Empena, Êpêra.

Two approaches to nasalization in Embera have been proposed: prosodic and segmental. Morris (1977) claims that nasalization is best analyzed as a prosodic, suprasegmental phenomenon that attaches to the syllable or word, applying to the whole prosodic unit at once and selecting the "nasal version" of all phonemes that can be nasalized. On the other

[6] See the phonemic inventory for the Embera of Antioquia and Córdoba in Pardo and Aguirre (1993:303).

hand, Loewen (1963a) and Rex and Schöttelndreyer (1972) consider nasalization to be segmental, and thus there are phonemically distinct nasal vowels that in turn affect other segments phonetically by spreading the nasal feature.

However, the abstract representation of the process of nasalization is irrelevant to this study. What is important for our purposes is that nasal spreading is a process that spreads the nasal feature rightward within a word, crossing affix boundaries and nasalizing all vowels and all consonants that allow it to continue (permeables, as Cohn calls them (1993:340)) until it reaches either a consonant that blocks nasalization (blockers) or the boundary of the phonological word. The case markers and pragmatic status markers discussed in this study are all subject to nasal spreading from the preceding word and so in that sense they are phonologically dependent on their host. This feature is discussed further in chapter 5.

As described in Cohn (1993:335), nasalized consonants such as we find in Katío emerge because of the spreading nasal feature. In Embera Katío, this spreading may stem from two sources: a nasal consonant as in (7a) or a nasal vowel as in (7b). As noted above, the nasal feature spreads rightward until it encounters either a phonological word boundary or a blocking consonant. At this juncture between a nasalized vowel and a blocking consonant, a transitional nasal homorganic to the following consonant is realized during the closure of the consonant's articulators. This transitional nasal occurs root-medially as in (7a) and also across affix boundaries as in (7b), but not across phonological word boundaries as in (7c).

(7) a. <*mebea*> [mẽ^{m}bea] 'sibling'
b. <*jũẽ-pe*> [hũẽ^{m}p^{h}e] 'after arriving'
c. <*mɯ papa*> [mũ̈ p^{h}ap^{h}a] / *[mũ̈^{m}p^{h}ap^{h}a] 'my mother'

We can divide the consonants of Katío into two groups: permeables and blockers. In the context of nasal spreading, sonorants[7] along with the implosive stops /ɓ/ and /ɗ/ are permeable to nasal spreading and are realized as their nasalized allophones. Elements to

[7] The element [h], which in some traditions is considered a glottal fricative, is classified as [+sonorant] in Chomsky and Halle (1968:303), and so on that basis I include it with the sonorants.

their right will also be affected by the nasal spreading. All other obstruents besides the implosives block nasal spreading and are realized with a homorganic transitional nasal.

In a nasal context, the third series of stops /ɓ/ and /ɗ/ become nasals. Distributionally, they cannot be found realized as implosive stops in nasal contexts, and are written as nasals in the orthography in such environments. Example (8) shows *b'* /ɓ/ 'AUX' first in an oral context in (8a), and then in a nasal context in (8b) where it is realized as [m]. Similarly, example (9) shows the locative *d'e* [ɗe] first in an oral context in (9a), and then in a nasal context in (9b), where it is realized as [nẽ].

(8) a. *k'o b'-ʉ-a*
[k^{h}o ɓɯa]
eat AUX-PRS-DECL
'eat/eats'

b. *k'o nʉ-b'-ʉ-a*
[k^{h}o nʊ̃mʊ̃ã]
eat PROG-AUX-PRS-DECL
'am/are/is eating'

(9) a. *do zake=d'e*
[do zak^{h}eɗe]
river small=LOC
'in the small river'

b. *do zroma=d'e*
[do zɾomãnẽ]
river big=LOC
'in the big river'

The nasalized allophones of the permeable consonants are shown with examples in table 3.

Table 3. Nasalized allophones of the consonants permeable to nasal spreading.

Nasalized Allophone	**Example**	**Gloss**
/ɓ/ [m]	/nɯ-ɓ-ɯ-a/ [nʊ̃mʊ̃ã]	'PROG-AUX-PRS-DECL' (is __-ing)
/ɗ/ [n]	/bania=ɗe/ [banĩãnẽ]	'water=LOC' (in water)
/ɾ/ [n̆] / [r̃] [nasal flap]	/mɯ=ra/ [mʊ̃n̆ã]	'1SG=TOP'
/w/ [w̃]	/mawu̯a/ [mãw̃ʊ̯̃ã]	'thus'
/j/ [ɲ] [palatal nasal]	/wã-ja/ [wãɲã]	'go-FUT'
/h/ [h̃]	/bãha/ [bãh̃ã]	'sky'

In the orthography, a vowel that follows a nasal consonant, such as *ʉ* in *mʉ* '1SG' or *e* in *mebea*, is nasalized by spreading from the preceding consonant, but is not written as nasal since the nasalization is predictable from context.

1.5.3 Introduction to Katío syntax

All the languages in the Chocó family have an ergative-absolutive alignment system (Mejía Fonnegra 2000:55). The ergative is marked with enclitic *ba* (*pa* in some languages) and the absolutive is unmarked, as discussed above in 1.2.

The basic word order of the Chocó languages is SOV (Aguirre Licht 2009:234). Example (10) illustrates the S, *mʉ mebea* 'my brother',[8] followed by the O argument *b'egi* 'deer' and finally V the verb *beasia* 'killed'.

(10) *Mʉ mebea =ba b'egi =ta bea-si-a.*
1SG sibling =ERG deer =FOC kill-PST-DECL
My brother killed a deer.

Although Katío is an ergative language, it shows nominative alignment in the verbal agreement system. Subject is a relevant category because the verbal morphology includes a plural subject agreement suffix *-d'a*. This affixes to transitive and intransitive verbs and thus indicates agreement with ergative subjects as in (11) and absolutive subjects as in (12).

(11) *Dai =ba ũrĩ-si-**d'a** k'orog'o trãbi =ta.*
1PL.EXCL =ERG hear-PST-PL snail traditional.song =FOC
We heard the traditional song of the snail. [EP:80]

(12) *Mobe wã-si-**d'a**-bida ãzhi.*
then go-PST-PL-EVID 3PL
Then they went (it is said). [FD:5]

In noun phrases, Embera Katío expresses a possessive or descriptive relationship through simple juxtaposition (Velupillai 2012:183). Possessors precede the head noun as in (13a), and adjectives follow the noun as in (13b).

[8] The words *mebea* 'sibling' and *kima* 'spouse' are ambiguous as to gender and may appear in the free translations as 'brother' or 'sister' and as 'husband' or 'wife' respectively, depending on the context.

(13) a. *mʉ zeze trʉ̃*
1SG father name
my father's name

b. *kucho zake*
knife small
small knife

Grammatical relations are coded for ergative, dative, and benefactive case with morphemes that occur after the arguments of the verb.

The ergative case marker *ba* marks the subject of a transitive verb. In (14) below, the name *Guillermo* is marked with *ba* because it is the subject of the transitive verb *k'õ* 'cut', while in (15) where it appears as the subject of the intransitive verb *trʉ̃a* 'sing' it is not marked with *ba*. The case marker *ba* is discussed in more detail in section 2.

(14) *Guillermo =ba b'ud'a =ta k'õ nʉ-m-ʉ-a.*
Guillermo =ERG hair =FOC cut PROG-AUX-PRS-DECL
Guillermo is cutting hair.

(15) *Guillermo trʉ̃ã nʉ-m-ʉ-a juwua ʉ̃ta DaiZeze =á.*
Guillermo sing PROG-AUX-PRS-DECL hand up God =DAT
Guillermo is singing with his hands up to God.

The dative case, or the third argument of a ditransitive verb in Embera Katío, is marked with *á*.[9] A dative noun phrase is semantically a recipient or addressee, according to Mortensen (1999:50). In example (15) above, *DaiZeze* 'God' is the goal of the verb *trʉ̃ã* 'sing' is in dative case. Example (16) shows the semantic recipient *mʉ* '1SG' in dative case.

(16) *Mʉ =á cobija =ta dea-tua.*
1SG =DAT blanket =FOC give-IMP
Give me the blanket. [EP:149]

The dative case marker *á* attaches to the right edge of a noun phrase, as in example (17) where its host is not the head noun, but rather the number *ʉ̃me* 'two'.

[9] Following Mortensen (1999), I use the accent mark to distinguish between the dative case *á* and the ergative allomorph *a* which is discussed in chapter 2, because their forms are indistinguishable in the case of those pronouns which take the allomorph *a*. In speech the pronunciation of *mʉa* '1SG=ERG' and *mʉá* '1SG=DAT', for example, are very similar, if not identical.

(17) *DaiZeze* =*ba* *ua* *dia-si-a* *zhumakĩrã-rã* *ũme* =***á***.
God =ERG garment give-PST-DECL man-PL two =DAT
God gave clothes to two men.

In example (18), the dative marker attaches to a nominal argument that includes a relative clause: *wũẽrã dai zeze bedea ĩjã bʉ* 'a woman who believed God's word'. Since the verb of the relative clause is the rightmost element in the phrase, *á* attaches to the verb.

(18) *Mawũã* *krĩcha* *b'-ʉ* =*d'e* *[wũẽrã* *DaiZeze* *bedea* *ĩjã* *b'-ʉ]* =***á***
do.thus think AUX-PRS =LOC woman God word believe AUX-PRS =DAT

bed'ea-si-a *DaiZeze* =*ta* *k'ãĩmok'ara* =*d'e*.
speak-PST-DECL God =FOC dream =LOC
As I am thinking like this, God spoke in a dream to a woman who believed his word.
[BBT:5]

The benefactive case marker *ita* expresses semantic roles such as beneficiary, recipient and goal. It is homophonous with the subordinating conjunction *ita* which marks certain purpose clauses. I gloss *ita* on noun phrases as 'BEN', following Mortensen (1999), and *ita* on purpose clauses as 'PURP'.

The benefactive case *ita* marks the beneficiary of an action in (19) and the goal of an action in (20).

(19) *Jozeka* =*ba* *zhiko* =*ta* *zhu-si-a* *Tito* =***ita***.
Jozeca =ERG food =FOC cook-PST-DECL Tito =BEN
Jozeca cooked food for Tito.

(20) *Dai* =*ba* =*ra* *kora-bodo* *nʉ-m-ʉ* =*ta* *tu-ba-d'a*
1PL.EXCL =ERG =TOP fall-about.to PROG-AUX-PRS =FOC cut-HAB-PL

tʉbʉ =***ita***.
firewood =BEN
We cut the ones [trees] about to fall for firewood. [V&G:p25]

Like the dative case marker *á*, the benefactive *ita* attaches not to the noun or pronoun itself but rather to the rightmost edge of the constituent or phrase. In (21), it occurs on the phrase *mʉ aba* 'me only'.

(21) *"Naũ bania =ra [mɯ aba] =**ita**", a-si.*
this water =TOP 1SG one =BEN say-PST
"This water is for me only," he said. [V&G:p89]

The morpheme *ita* is one of several purpose clause subordinators in Embera Katío, and in this context I gloss it 'PURP'. In example (22), *ita* marks subordinate clauses "for where to make the holes," and "to bury the posts." In example (23), *ita* marks the purposive clause "so we won't be sick."

(22) *Mobe-bɯrɯ nara za b'-a-ri-a [zobea sãma o-koa]*
then-COND first measure AUX-NPRS-HAB.PRS-DECL hole where make-REPT

*=**ita**, [drũdaɯ jou-kua] =**ita**.*
=PURP house.pole bury-PL =PURP
Then first it is measured for where to make the holes, and to bury the house posts.
[V&G:p65]

(23) *Ẽbẽra =ba k'ore zhara k'o-d'a-ka, mawuamina izhi ɯmɯ*
person =ERG crocodile meat eat-PL-NEG.HAB however 3SG egg

k'o-ba-d'a, [dazhi k'aya ẽ b'-a] =ita.
eat-HAB-PL 1PL.INCL sick NEG AUX-NPRS =PURP
Embera don't eat crocodile meat, but they eat its eggs, so [we won't be sick].
[V&G:p126]

Example (24) displays all three case markers in one utterance. The argument *zhuwũẽrã* 'woman' is coded for ergative case, absolutive *zhiko* 'food' is unmarked for case and coded as topical, *jõma* 'all' is coded for benefactive case, and three arguments are coded for dative case: *warrarã* 'children', *zhi kimá* 'her husband', and *wabemarã* 'others'.

(24) *Zhuwũẽrã =ba zhiko =ra epedeko =d'e jõma =ita tue-ya: zhi*
woman =ERG food =TOP plate =LOC all =BEN serve.food-FUT 3SG

warra-rã =á, zhi kima =á, wabema-rã =á sida...
son-PL =DAT 3SG spouse =DAT others-PL =DAT also
The woman will serve the food on a plate for everyone: to her children, to her husband, and also to others... [V&G:p37]

The examples in this section provide a very basic background for understanding the grammar of Katío and the examples in this thesis. In the following chapters I focus on

three morphemes. One of these is a case marker indicating grammatical relations: the ergative case marker *ba*. The other two mark pragmatic status: *ta* marks focus and *ra* marks topic. Finally in chapter 5, I present a summary of the syntactic and phonological features of these morphemes that are relevant to their classification as words, affixes or clitics.

CHAPTER 2

ERGATIVE *BA*

All of the researchers introduced in section 1.2 discuss *ba* or its equivalent. Loewen (1958:92, 99) describes *-pa* as an "agentive" inflectional suffix, marking "subject-agents" and instruments, as well as another *-pa* or *-upa* ablative suffix of source translated as "from".[1] Schöttelndreyer (1977:147) also considers it a suffix of "active subject" and mentions in an endnote that it is "a case marker that indicates that the agent acts as an instrument." Rex (1975:38) calls *ba* an enclitic ergative case marker that marks the subject of a transitive clause, instrument, and "locative of source," connected by the "semantic notion of source." According to Harms (1994:10), Epena Pedee (SE) has an ergative case suffix *-pa* with an allomorph *-a* with selected pronouns. He also describes a "cause suffix" *-pa*, which "marks origin" and which he glosses as "from" in his examples (1994:13), but which he does not seem to connect with ergative *-pa*. Aguirre Licht (1999:319) also describes an ergative suffix *-ba* with an allomorph *-a* in Embera Chamí (SE). Like Rex, Mortensen (1999:47) recognizes several uses for *ba* including ergative case, instrument and reason, which he connects with the sense of "cause or origin" and labels "ablative."

I interpret Loewen and Schöttelndreyer's descriptions of *ba* as "subject-agent" and "active subject" to be attempts to describe the notion of ergative case. If this is so, then all the researchers are essentially in agreement that *ba* marks ergative case. Four also mention that it marks instrument.

The disagreement thus centers around its status as a suffix or enclitic and whether it also marks source as according to Rex or reason/cause as according to Mortensen and Harms. The question also remains as to whether these instances of *ba* are better understood as different usages of a single morpheme as Rex and Mortensen suggest or as different morphemes, as Harms suggests.

[1] See question for further research about ablative *ʉba* 'from' in 6.3.

In my data corpus I found instances of *ba* marking ergative case, instrument and reason, shown below with examples. I did not find any clear instances of *ba* expressing source, and found that the examples of *ba* as source given by Rex (1975:39) could also be interpreted as instances of the postposition *u̶ba* 'from'. The question of whether this and other postpositions of source may have a relation to the ergative *ba* is one for future research and is discussed further in section 6.3.

2.1 The three functions of *ba*

In this section, I present examples of the different functions of *ba*, marking ergative case, instrumental case and reason. Then in section 2.2, I discuss the issue with applying a broad term such as "agentive" or "ablative" to generalize all functions of *ba*.

Ergative characteristics are found at some level in roughly one fourth of the world's languages (Dixon 1994:2). Languages that have ergative alignment in their basic grammatical relations treat the subject of an intransitive verb (S) and the object of a transitive verb (P) the same (absolutive case), and the subject of the transitive verb (A) differently (ergative case) (Payne 1997:135).

In Embera Katío, ergative case is marked with *ba*. Example (25) shows the simple ergative function of *ba*, marking *ẽbẽra* 'Embera person' as the subject of the transitive verb *biradu̶ga* 'step on'. In contrast, (26) demonstrates that *zhumakĩrã* 'man', the subject of the intransitive verb *jeweda* 'sit', is unmarked, and (27) shows that *zhumakĩrã* is ungrammatical when marked with *ba*. The ergative case marker *ba* may only mark a subject of a transitive verb.

(25) *Ẽbẽra* =***ba*** *kẽberre* =*ta* *biradu̶ga-ba-ri-a.*
person =ERG cockroach =FOC step.on-HAB-HAB.PRS-DECL
The Embera steps on cockroaches.[2] [V&G:p118]

(26) *Zhumakĩrã* *jeweda-si-a.*
man sit-PST-DECL
The man sat.

[2]See footnote 3 in chapter 2.

(27) * *Zhumakĩrã* =*ba* *jeweda-si-a.*
man =ERG sit-PST-DECL
The man sat.

The ergative case marker may occur on human arguments such as *ẽbẽra* 'Embera' in (25) above, non-human animate arguments such as *bid'o* 'wild boar' in (28) below, and inanimate arguments such as *bisia* 'sunlight' (not shown here), and *bakuru* 'tree' in (29).

(28) *Bid'o* =*ba* *ne* *jõma* *k'o* *chu* *b'-ʉ-a.*
wild.boar =ERG thing all eat STV be-PRS-DECL
The wild boar eats everything.[3] [V&G:p39]

(29) *Bakuru* =*ba* =*ra* *neta* *dia* *nʉ-m-ʉ-a.*
tree =ERG =TOP fruit give PROG-AUX-PRS-DECL
The tree is giving fruit.

As described by Harms and Aguirre Licht regarding Southern Embera languages, in Katío the ergative enclitic *ba* has an allomorph *a* that occurs on most personal pronouns, with the exception of 1PL exclusive *dai* and 2PL *mãrã*.

Table 4. Ergative case personal pronouns

	Singular	**Plural**
1st Person	*mʉ*=*a*	*dai*=*ba* (EXCL)
		dazhi=*a* (INCL)[4]
2nd Person	*bu*=*a*	*marã*=*ba*
3rd Person	*izhi*=*a*	*ãzhi*=*a*[5]

The allomorph *a* only occurs on these personal pronouns; all other words take the form *ba*.

Although the great majority of ergative arguments are marked with *ba*, it is occasionally not present on an ergative case element in contexts where there is no chance of

[3] In Katío, generic subjects are often singular while in English a generic subject is more often expressed as plural. Therefore, a more equivalent translation for (25) in English would be "Embera people step on cockroaches," and for (28), "Wild boars eat everything." Since this does not reflect the morphology of the example, however, I have translated these as singular in the free translations.

[4] There exists also a redundantly plural alternative form *dazhi-rã*=*ba* '1PL.INCL-PL=ERG'

[5] In pronouns ending in i that take allomorph =*a*, the -i is often omitted in the phonetic realization. I.e. *dazhi*=*a* ['dadʒa], *izhi*=*a* ['idʒa], *ãzhi*=*a* ['ãⁿdʒa]

misinterpretation of the ergative argument's grammatical relation to the verb, and its absence does not appear to affect the meaning. Example (30) shows two grammatical ways of asking "what are you (pl.) eating?" The first has ergative marker *ba*, the second has the unmarked pronoun.[6]

(30) a. *Marã* =***ba*** *k'ãrẽ* =*ta* *k'o* *nu̶-m-u̶-pe?*
2PL =ERG what =FOC eat PROG-AUX-PRS-Q
What are you (pl.) eating?

b. *Marã* *k'ãrẽ* *k'o* *nu̶-m-u̶-pe?*
2PL what eat PROG-AUX-PRS-Q
What are you (pl.) eating?

Similarly, in (31), only the instrument *juwua* 'hand' is marked with *ba*, while the animate agent *wũẽra zakerã* 'girls', which we would also expect to be marked with *ba* is not marked with the ergative case marker. It also happens to have the topic marker *ra*, discussed in 4, but note that the fact that the ergative argument is marked as topical does not mean it cannot also be coded for ergative case. In (32), both ergative *ba* and topic *ra* are present.

(31) *Wũẽrã* *zake-rã* =*ra* *juwua* =***ba*** *b'ed'a* *bea* *pan-a-ri-a...*
woman small-PL =TOP hand =INS fish kill AUX.PL-NPRS-HAB.PRS-DECL
As for girls, they kill fish with their hands... [V&G:p32]

(32) *Warru* =***ba*** =***ra*** *eterre* =*ta* *k'o* *b'-a-ri-a.*
panther =ERG =TOP chicken =FOC eat AUX-NPRS-HAB.PRS-DECL
The panther eats chickens. [V&G:p196]

Instrumental case is also marked with *ba*. It is typologically common for instrumental case and ergative case to be identical in form. In fact, more than three quarters of Australian languages have one form for both ergative and instrumental case (Dixon 2011:304).

In example (33), the animate agent *ẽbẽra* 'person' is marked with ergative *ba* and the instruments *duga* 'hook' and *miasu* 'spear' are marked with instrumental *ba*.

[6] Note that much remains for future study such as classes and valence of verbs and how that may affect the coding of the arguments.

(33) a. *Ẽbẽra* =*ba* *bea-ba-ri-a* *duga* =***ba,***
person =ERG kill-HAB-HAB.PRS-DECL hook =INS
An Embera kills with a hook,

b. *miasu* =***ba*** *bea-ba-d'a.*
spear =INS kill-HAB-PL
with a spear they kill. [V&G:p43]

In example (34), the instrument used to kick, *jĩrũ* 'foot', is marked with *ba*.

(34) *Jĩrũ* =*ba* *tɯga-koa-pe,* *Pirapoi-tua!*
foot =INS kick-REPT-SEQ[7] get.up-IMP
He kicked them with his foot, "Get up!" [EP:72]

The enclitic *ba* also marks reason. In (35), because the verb *wã* is intransitive, the agent *mɯ* '1SG' is an absolutive and is marked with topic marker *ra* (section 4). However, *maɯ̃* 'this', which refers to preceding events in the text, is marked with *ba*, expressing reason.

(35) *Mɯ* =*ra* *maɯ̃* =***ba*** *wã* *ẽ-b'-a-si-a.*
1SG =TOP that =REAS go NEG-AUX-NPRS-PST-DECL
Because of this I didn't go. [Africanas:2]

A single-word explanation or answer to a question may also be expressed with the reason marker *ba*, as in (36b).

(36) a. *K'ãrẽã* *jẽga* *nɯ-m-ɯ-pe?*
why cry PROG-AUX-PRS-Q
Why [are you] crying?

b. *Sopua* =***ba.***
sad =REAS
Because [I'm] sad.

Of course, ergative *ba* and reason *ba* may co-occur in the same sentence, and are both present in example (37).

[7] The repetitive suffix *-koa* or *-kua* is also sometimes used to indicate a plural object, as in (34).

(37) *Maʉ̃* =***ba*** *dai* =***ba*** *krĩcha* *pan-ʉ-a* *nusi* =*ra* *do*
this =REAS 1PL.EXCL =ERG think AUX.PL-PRS-DECL wolf.fish =TOP river

papa =*ta.*
mama =FOC
Because of this we think that the wolf fish is the river's mother. [V&G:p151]

The morpheme *ba* behaves the same syntactically for all three functions. It attaches to the rightmost edge of a constituent, and so it appears on adjectives or quantifiers modifying the head noun, or on a complement clause. In example (38) ergative case *ba* is attached to the adjective *zake* 'small' in the noun phrase *wũẽrã zake* 'small woman' or 'girl'.

(38) *[Wũẽrã zake]* =***ba*** *yuk'a* =*ta* *ẽ* *nʉ-m-ʉ-a.*
woman small =ERG yucca =FOC dig PROG-AUX-PRS-DECL
The young woman is digging yucca.

In example (39), *ba* attaches to the quantifier *ume* 'two'.

(39) *De* *o-d'a-i* =*ra* *nara* *akʉ-d'a-i* *bara* *ẽuja* =*ta* *[zhi* *kima* *ũme]*
house do-PL-IRR =TOP first look-PL-IRR OBLG land =FOC DEF spouse two

=***ba.***
=ERG
As for making a house, first the two spouses must look at the land. [V&G:p56]

Example (40) shows the instrumental case *ba* attached to the rightmost edge of the phrase *duga zroma* 'a big hook'.

(40) *Ãparra* *[duga* *zroma]* =***ba*** *bea-ba-ri-a.*
sp.of.fish hook big =INS kill-HAB-HAB.PRS-DECL
The fish is killed with a big hook...[8] [V&G:p20]

The enclitic may also occur on the rightmost edge of a complement clause, attaching to the verb, as in (41) where reason *ba* occurs on the inflected verb *ũdu* 'see'.

[8] Embera Katío has no voice-alternation (passivizing) morphology, but a passive voice is achieved by simply omitting the ergative argument, as in example (40). Compare examples (40) and (33).

(41) *Mawũẽ [mɯ maã̄ plata ũdu-bɯrɯ]* =***ba*** *sobia nɯ-m-ɯ* =*ne ed'a*
then 1SG this money see-PRS =REAS happy PROG-be-PRS =LOC LOC

wetá-bɯrɯ sobia-si-a.
more-EMPH happy-PST-DECL
Then being happy because I saw this money, I became even happier. [BBT:122]

2.2 Agentivity

If ergative case, instrumental and reason are all marked with the same form *ba*, we must consider whether they might be a single morpheme rather than three. Reesink offers a guideline for analyzing identically shaped morphemes: "One form stands for one meaning, unless the reasoning necessary to arrive at the one meaning seems too much contrived" (1981:91). What meaning can we identify to unite these three forms?

A possible term to encompass ergative, instrumental and reason is that of agentivity. Agentivity may be defined in terms of Grimm's semantic lattice made up of four properties: volition, sentience, instigation and motion (2005:21) which are based on Dowty's definition of the Agent Proto-role (1991:572). According to Butt, ergative case is the most prototypically agentive in that it may be associated with all four properties or just volition and sentience. Instrumental may entail both instigation and motion or only one of these two (2006:86). Reason must include at least instigation of an affected change or state. These properties of the agentive proto-role suggest a foundation to group these three functions of *ba* together semantically since each function must entail at least one of Grimm's four agentive characteristics.

Loewen (1958:99) uses the term agentive, and it may at first appear to neatly tie together the three uses of *ba*. However, we must note that agentive is a semantic term and ergative case is a purely syntactic phenomenon that, although often corresponding to the semantic agent, is not strictly agentive. For example, in Katío, verbs like *ũrĩ* 'hear' in (42), *ũdu* 'see' (43b), and *k'awua* 'know' in (44), take subjects marked for ergative case, but in terms of semantic roles, their subjects are not agents but rather experiencers.

(42) *Dai* =***ba*** ***ũrĩ****-si-d'a k'orog'o trãbi* =*ta.*
1PL.EXCL =ERG hear-PST-PL snail traditional.song =FOC
We heard the traditional song of the snail. [V&G:p127]

(43) a. *Kaĩ ẽbẽra =ba mɯ zeze =ra bea-si-a.*
that person =ERG 1SG father =TOP kill-PST-DECL
That person killed my father.

b. *Mɯ kima =**ba** **ũdu**-si-a.*
1SG spouse =ERG see-PST-DECL
My wife saw [him].

(44) *K'ai =**ba** **k'awua**-i-pe?*
who =ERG know-IRR-Q
Who knows? [EP:93]

Thus we cannot assume a perfect correlation between semantic agent and ergative case, however tempting it may be.

Mortensen (1999:47) summarizes the three functions of *-ba* (*-pa* in Northern Embera) as expressing "cause or origin", or "ablative", which can refer to a case that marks instrumental (Crystal 1997) or source. Rex's unification of the three functions is similar in that they "have to do with the semantic notion of source" (1975:38). However, even were we to include the notion of source in the functions of *ba* (which I have not found evidence for in my data, as noted above), these terms are not sufficient to encompass those same instances of ergative case *ba* in Katío mentioned above that are problematic for the notion of agentivity. For example, the ergative case subject of the verb *ũrĩ* 'hear' cannot be understood as cause, origin, an instrument or a semantic source, and therefore neither ablative nor source is an adequate term to encompass all instances of *ba*.

I suggest, then, that while the notion of agentivity may be the most compelling and may give us a possible historical motivation for the three functions of *ba*, it cannot accurately describe all of its ergative instances. I recognize the connection between the three functions of *ba* and the semantic motivation to class them together, but I also recognize the shortcomings of the notion of agentivity to completely encompass the ergative case.

For describing Embera Katío, I find it useful to acknowledge the semantic similarities between *ba*'s three functions, but not to insist on a term that can encompass all three without exception, instead glossing each distinctly according to its function.

CHAPTER 3

FOCUS MARKER *TA*

The second morpheme I examine in this study is *ta*. It was briefly introduced in section 1.2 and discussed in relation to absolutive case in 1.3, and I begin by describing in more detail what other researchers have written in section 3.1, and the notion of focus in section 3.2. Section 3.3 shows the absolutive distribution of *ta* and section 3.4 shows examples of *ta* on focused complement clauses. Finally, section 3.5 discusses *ta* in comparison to the marked narrow focus of *tru* and *buru*.

3.1 Previous analyses of *ta*

Loewen describes *ta* as an "extensively used... enclitic emphatic particle" (1958:101), but makes no mention of absolutive case. Schöttelndreyer calls *ta* a "conflict suffix" and mentions that its presence "implies an impending conflict" (1976:254, 250).

Rex briefly discusses an enclitic *ta*, and uses the terminology "nominative," which apparently is meant to be equivalent to absolutive case as it marks "object of a transitive clause" and the "subject of an intransitive clause." She also shows examples of absolutives marked by *ra* because they are "previously mentioned" (1975:39-40).

Harms cites Rex's analysis in a footnote about Katío (1994:65), but in his own work on Epena Pedee, he describes a focus enclitic *ta* used "to focus on or draw attention to the clausal constituent to which it is attached" (1994:193).

Mortensen, analyzing Northern Embera, calls *ta* an absolutive focus suffix which marks "introductory focus" or focus on an absolutive that is "introduced in a discourse or recalled after being in the background" (1999:143).

The Waunana language, which is not an Embera language but is still within the Chocó family, also has a morpheme *ta*, described by Sánchez et al. in their pedagogical grammar

as an "emphatic particle" which occurs on "the most important word of the phrase that expresses the new or key information of a question or situation" (1977:100).

I argued in section 1.2 that *ta* is not an absolutive case marker, but instead marks the pragmatic status of focus on a constituent, as suggested by the terminology used by Harms (1994) and Mortensen (1999). This function is probably analogous to the description of *ta* given by Sanchez et al. (1977), and is perhaps also what Loewen (1958:101) tried to describe as an "emphatic particle."

The next section, 3.2, discusses of the concept of focus to lay the groundwork necessary for an adequate description of the function of *ta*.

3.2 Notion of focus

The term focus has been used liberally to refer to various linguistic phenomena and so requires careful defining. For the purpose of this thesis, focus is a pragmatic status. A pragmatic status "relates the content to the context" (Payne 1997:261). A pragmatic status marker guides the listener's handling of the flow of information, signaling how to relate a particular information component to its context.

According to Dooley and Levinsohn, focus refers to the part of an utterance that is most salient: the information that marks a change in the hearer's mental representation. They describe it as follows: "Material in focus typically (1) adds new information or (2) changes what is already present in an activated propositional framework... Every utterance has a focus" (2001:62). Lambrecht explains, "...The focus of the proposition expressed by a sentence in a given utterance context, is seen as the element of information whereby the presupposition and the assertion DIFFER from each other. The focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the UNPREDICTABLE or pragmatically NON-RECOVERABLE element in an utterance" [emphasis in original] (Lambrecht 1994:207, cited in Van Valin and LaPolla 1997).

The distinction that Gundel and Fretheim (2004) make between referential and relational givenness-newness is important to the discussion of *ta* and the definition of focus as I use it. Referential givenness-newness refers to "a relation between a linguistic expression and a corresponding non-linguistic entity in the speaker/hearer's mind." It refers to

the status of the *referent* of a linguistic element in the mind of the speaker and hearer, and includes concepts such as referentiality, specificity, familiarity, identifiability, and discourse activation status (Gundel & Fretheim 2004:176). These concepts are properties of the speaker/hearer's mental representation regarding the referent.

Relational givenness-newness, on the other hand, "reflects how the informational content of a particular event or state of affairs expressed by a sentence is represented and how its truth value is to be assessed" (2004:176). It refers to the relation of one part of the conceptual representation to the other and includes concepts such as topic, comment, presupposition, focus, and predicate. These concepts are properties only of linguistic representations (Gundel & Fretheim 2004:176) and do not connect to real-world referents. We might think of it as referring to the difference between the mental representation of the hearer before and after the utterance.

The relational givenness-newness concepts of topic and focus correlate with referential givenness-newness properties, and this is why they are easily confused. For example, topics are almost exclusively familiar and identifiable (Gundel 1988:212-214). They are also frequently activated in the discourse. However, this does not mean that a topic marker is therefore coding familiarity, identifiability or activation status as such. In the same way, a focus is often new and previously unactivated, but this does not mean that a focus marker is therefore coding a referent's introduction to the discourse. The terms topic and focus as used in this thesis refer to relational givenness-newness representations, and not to the referential properties usually associated with them.

Mortensen's description of *ta* is that it marks "introductory focus" on an absolutive that is "introduced in a discourse or recalled after being in the background" (1999:143). What he is describing with the term focus is the referential property of activation status, not the focus as used by Gundel and Fretheim (2004) and in this thesis.

In Katío, *ta* does not only occur in narrative texts with participants or entities to be activated and recalled, but it also occurs frequently in elicited data and isolated sentences that are not tracking participants. This leads me to believe that its function has less to do with the activation status of entities (the referential givenness-newness), and more to do

with the focus structure of the sentence and the fact that the information marks a change in the hearer's mental representation (the relational givenness-newness).

Examining the notion of focus in more detail, we find that focus varies in scope depending on the context of the utterance and how much of the utterance is focused information. Lambrecht differentiates between narrow and broad focus (1994:296). Narrow focus consists of a single argument, word or even morpheme (e.g. "I didn't say it was small, I said it was small*ISH*"). Broad focus includes more than a single constituent: It may be predicate focus ("That house I told you about, *THEY ALREADY SOLD IT*") or sentence focus (e.g. "*THERE ONCE WAS A GIRL FROM KENTUCKY*").

Since *ta* occurs on single arguments and complement clauses and not on predicates or sentences, for the purposes of this study I am more interested in how Embera Katío expresses narrow focus and the function that *ta* plays.

Two types of narrow focus can be distinguished: unmarked and marked (Lambrecht 1994:296). As described in Van Valin and LaPolla (1997:209), the difference depends on the position of the argument: whether it falls inside or outside of a language's default focus position.

Looking at focus patterns cross-linguistically, there is no universal position where focused information is presented (Harris 2002:45). Instead, the default focus position is language-specific: every language has a position in which focused information is typically presented. A focused constituent that appears in this default focus position is in unmarked narrow focus, while a focused constituent that occurs outside of this default focus position is in marked narrow focus (Van Valin & LaPolla 1997:209).

For example, in English the default focus position in a statement is core-final (Van Valin & LaPolla 1997:209). In the sentence "John gave a sandwich to Tim", the argument "to Tim" is in the default focus position and automatically receives unmarked narrow focus. It is ambiguous whether or not "to Tim" is in narrow focus or whether the utterance has predicate focus. In other words, the utterance "John gave a sandwich to Tim" may serve as the answer to two distinct questions: "Who did John give a sandwich to?", resulting in narrow focus on "to Tim" and "What did John do?", resulting in predicate focus on "gave a sandwich to Tim."

In English, in order to put marked narrow focus on an argument that is outside of the unmarked focus position, that argument is highlighted with high-pitch intonation: "JOHN gave a sandwich to Tim" or "John gave a SANDWICH to Tim" (Van Valin & LaPolla 1997:209-210).

In order to discover the default focus position for Embera Katío and explore the system for narrow focus, I examine the structure of content questions and answers. According to Harris (2002:46), the question word in a content question (sometimes called a WH question) and the element in the response that answers it are always in narrow focus.

In verb-final languages, the default focus position is normally the immediately pre-verbal slot (Van Valin & LaPolla (1997:209) cite Kim 1988), and this is the case in Embera Katío. This is shown in examples (45)-(48), in which the pre-verbal slot is preferred for the focused content question words, whether it be questioning P, S, A or an oblique. The focused question words are underlined.

(45) *ta* on P argument *k'ãrẽ* 'what'
Daizeze =*ba* *bu* =*á* =*ra* *k'ãrẽ* =*ta* *jara-si.*
God =ERG 2SG =DAT =TOP what =FOC say-PST
What did God say to you?

(46) *ta* on S argument *k'ai* 'who'
K'ai =*ta* *jũẽ-si* *puwuru* *ed'a* =*ra?*
who =FOC arrive-PST town LOC =TOP
Who arrived in the town?

(47) No *ta* on A argument *k'ai* 'who'
Usa *k'ai* =*ba* *bea-si?*
dog who =ERG kill-PST
Who killed the dog?

(48) No *ta* on oblique argument *sama* 'where'.
Bu *mebea* *sãma* *wã?*
2SG sibling where go
Where is your brother going?

The default focus position in Embera Katío is immediately pre-verbal. An argument in this slot will be in unmarked narrow focus.

Note that only those question words that are absolutives are marked with *ta*, in examples (45) and (46). Ergative *k'ai* 'who' in example (47) and oblique *sãma* 'where' in example (48) are also in narrow focus, but are not marked with *ta*. This distribution is discussed in the following section.

3.3 Distribution of *ta*

The focus marker *ta* is often optional. It most frequently marks the object as in example (49a), where it marks *pada* 'plantain', but it can also mark the subject of an intransitive verb as in example (50a). Examples (49b) and (50b) with absolutives unmarked with *ta* are also grammatical.

(49) a. *Mʉ pada =ta wia nʉ-m-ʉ-a.*
1SG plantain =FOC cook PROG-AUX-PRS-DECL

b. *Mʉ pada wia nʉ-m-ʉ-a.*
1SG plantain cook PROG-AUX-PRS-DECL
I am cooking plantain.

(50) a. *Ewari aba mʉ mebea =ta jũẽ-si-a puwʉrʉ ed'a.*
day one 1SG sibling =FOC arrive-PST-DECL town LOC

b. *Ewari aba mʉ mebea jũẽ-si-a puwʉrʉ ed'a.*
day one 1SG sibling arrive-PST-DECL town LOC
One day my brother arrived in the town.

The focus marker *ta* marks narrow focus on absolutive constituents. In order to substantiate this claim, I again refer to the content questions that reveal the default focus position. When an absolutive element is questioned, the answering element is usually in the unmarked focus position and marked with *ta*.

Example (51) is a simple statement, independent of a question. In example (52), (52a) is a question, and (52b) is the answer to the question.

(51) *Mʉ pada k'o nʉ-m-ʉ-a.*
1SG plantain eat PROG-AUX-PRS-DECL
I'm eating plantain.

(52) a. *Bu* =*ra* *k'ãrẽ* =***ta*** *k'o* *nʉ-m-ʉ-pe?*
2SG =TOP what =FOC eat PROG-AUX-PRS-Q
What are you eating?

b. *Mʉ* =*ra* *pada* =***ta*** *k'o* *nʉ-m-ʉ-a.*
1SG =TOP plantain =FOC eat PROG-AUX-PRS-DECL
I'm eating plantain.

Note how in the simple statement (51), both arguments may be unmarked, as there is no need to distinguish between the ergative and absolutive arguments since it is unlikely to be interpreted as "plantain is eating me."

In contrast, the marking of the arguments in the answer (52b) should be parallel to that of the question (52a). In (52b), *mʉ* takes *ra*, echoing *bu*=*ra* in (52a), and *pada* takes *ta*, echoing *k'ãrẽ*=*ta*. Thus (51) is grammatical on its own, but not as a response to (52a).

The *ta* morpheme is not necessary (but is acceptable) in a simple statement such as (51). It performs a function that is related to the information focus structure that we find in questions. Note that the two constituents marked with *ta* in (52a) and (52b) are both in the pre-verbal default focus position.

In example (52) above, it is the P argument that is questioned. The S argument may also be questioned and marked with *ta*, as in (53), where *k'ai* 'who' is marked with *ta*.

(53) *K'ai* =***ta*** *jũẽ-si* *puwʉrʉ* *ed'a* =*ra?*
who =FOC arrive-PST town LOC =TOP
Who arrived in the town?

Two acceptable answers to the question "who arrived in the town?" are shown in examples (54) and (55). In both, the answering constituent is marked with *ta*, echoing the questioning constituent *k'ai*=*ta* 'who=FOC'. In the first, *mʉ mebea*=*ta* is not in the pre-verbal slot, and in the second, *mʉ*=*ta* '1SG-FOC' is immediately pre-verbal and occupies the default focus position. Focused constituents marked with *ta* may occur in or outside of the default focus position.

(54) *Mʉ* *mebea* =***ta*** *puwʉrʉ* *ed'a* *jũẽ-si-a.*
1SG sibling FOC town =LOC arrive-PST-DECL
My brother arrived in the town.

(55) *Mʉ* =***ta*** *jũẽ-si-a* *puwʉrʉ* *ed'a* =*ra.*
1SG =FOC arrive-PST-DECL town LOC =TOP
I arrived in the town.

Both examples (54) and (55) are grammatical with either *mʉ mebea* '1SG brother' or *mʉ* '1SG' as the subject. The difference in word order is due to the postposing of the postpositional phrase[1] *puwʉrʉ ed'a* 'in the town' in order to allow *mʉ mebea* to be in the focus position. The process of postposing is discussed in section 4.3.

In the above two pairs of question and answer ((52) and (53) answered in (54)/(55)), the questioned constituent was an absolutive. I now examine a case in example (56) in which an ergative constituent is questioned, and focus morpheme *ta* is not used. The focused question constituent *k'ai* 'who' is marked only with ergative *ba*. Notice that the word order of this question is OSV, illustrating the variety of permissible word orders in Katío. This word order is motivated by the preference for the focused question word *k'ai* 'who' to be in the pre-verbal default focus position.

(56) *Usa* *k'ai* =*ba* *bea-si?*
dog who =ERG kill-PST
Who killed the dog?

Another possible phrasing of the same question is shown in example (57), and keeps the questioned constituent *k'ai* 'who' in the pre-verbal focus position, but moves the P argument *usa* 'dog' to follow the verb. Example (58) is identical except *usa* is unmarked.

(57) *K'ai* =*ba* *bea-si* *usa* =*ra?*
who =ERG kill-PST dog =TOP
Who killed the dog?

(58) *K'ai* =*ba* *bea-si* *usa?*
who =ERG kill-PST dog
Who killed the dog?

[1] The postpositional phrases in this paper could also potentially be analyzed as noun phrases with locative case markers. Some of the postpositions show similar behavior to the case markers, shifting the stress of the preceding word and being subject to nasal spreading, as discussed in chapter 5, but I do not have enough data for most of the postpositions to draw a conclusion about their class. Therefore, for the purposes of this thesis, I will call them postpositions to distinguish them from case markers, but this distinction is somewhat arbitrary.

Example (59) shows that it is ungrammatical to attach *ta* to a focused ergative constituent with *ba*.

(59) * *K'ai* =***ba*** =***ta*** *bea-si* *usa* =*ra?*
who =ERG =FOC kill-PST dog =TOP
Who killed the dog?

However, if we attach *ta* to the same constituent without *ba*, it forces an interpretation of the sentence so that the argument may be understood as an absolutive. In example (57) above, the ergative marker *ba* on the interrogative pronoun *k'ai* 'who' indicates that the questioned constituent is the agent of the verb, requiring an interpretation of 'who killed the dog?', while in example (60) below, the focus marker *ta* indicates that the questioned constituent is the focused patient of the verb requiring an interpretation of 'who did the dog kill?'

(60) *K'ai* =*ta* *bea-si* *usa* =*ra?*
who =FOC kill-PST dog =TOP
Who did the dog kill?

Therefore, if a focused constituent is an ergative, as in (57), it cannot be marked with *ta*, whereas if it is an absolutive it is marked with *ta*.

As we would expect of a focus marker, it is ungrammatical to mark an absolutive argument that is not in focus with *ta*. In (61), the ergative *k'ai* 'who' is in focus. If *ta* were simply an absolutive case marker, as posited by Rex (1975) and Schöttelndreyer (1977), we would expect example (61) to be grammatical, since *usa* 'dog' is an absolutive. However, since *ta* is not an absolutive marker, but rather a focus marker, and in this case *usa* is not in focus, it must be marked with *ra* as in (57) or unmarked as in (58).

(61) * *K'ai* =*ba* *bea-si* *usa* =*ta?*
who =ERG kill-PST dog =FOC
Who killed the dog?

Example (62) is an acceptable answer to the question "who killed the dog?" in (57). The answering constituent *imama* 'tiger' is immediately pre-verbal in the default focus

position, and the other argument is postposed and marked with topic marker *ra* (see 4.3), for a word order of SVO. This order mirrors the order of the question in (57).

(62) *Imama =ba bea-si usa =ra.*
jaguar =ERG kill-PST dog =TOP
The jaguar killed the dog.

I now examine cases in which it is an oblique of direction, location or time which is questioned. In (63), the questioned constituent is an oblique of direction *sãma* 'where' which falls in the pre-verbal default focus position and is unmarked.

(63) *Bu mebea sãma wã?*
2SG sibling where go
Where is your brother going?

A question word such as *sãma* 'where' that is not an absolutive can not be marked with *ta*, as shown in example (64).

(64) * *Bu mebea sãma =ta wã?*
2SG sibling where =FOC go
Where is your brother going?

Since an ergative argument or an oblique cannot take the focus marker, the only way to mark either of them as being in narrow focus is by placement in the default focus position. It follows that the focused question word will be ungrammatical if postposed, as in (65), or if fronted, as in (66).

(65) * *Bu mebea wã sãma?*
2SG sibling go where
Where is your brother going?

(66) * *Sãma bu mebea wã?*
where 2SG sibling go
Where is your brother going?

Like in example (61) above, contrary to what we would expect if *ta* were a simple absolutive case marker, it is ungrammatical to attach *ta* to an absolutive case constituent that is not in focus, so example (67) is not acceptable.

(67) * *Bu mebea =ta sãma wã?*
2SG sibling =FOC where go
Where is your brother going?

In the response to the question "where is your brother going?" in (68), the answering element also falls in the pre-verbal slot. And once again, it is ungrammatical for an absolutive element that is not in focus to take *ta*, shown by example (69).

(68) *Mʉ mebea do ed'a wã.*
1SG sibling river LOC go
My brother goes to the river.

(69) * *Mʉ mebea =ta do ed'a wã.*
1SG sibling =FOC river LOC go
My brother goes to the river.

Obliques of location are similar: the focused oblique question word will fall in the pre-verbal default focus position, but it will be unmarked, as in (70). Because it is not an absolutive, the oblique answering constituent does not take *ta*, shown in (71). Even though the postpositional phrase *de edre* 'under the house' is in narrow focus, it is ungrammatical for it to take focus enclitic *ta*, as shown in (72).

(70) *Mog'ara sãma chu-pe?*
rock where STV-Q
Where is the rock?

(71) *Mog'ara =ra de edre chu-a.*
rock =TOP house under STV-DECL
The rock is under the house.

(72) * *Mog'ara =ra de edre =**ta** chu-a.*
rock =TOP house under =FOC STV-DECL
The rock is under the house.

Just like with ergative arguments and other obliques of direction and location, *ta* cannot attach to obliques of time. Example (73) shows the correct question and answer.

However, the answer in example (74) is ungrammatical with *ta* attached to the oblique of time *nu* 'tomorrow', even though it is in focus because it answers the preceding question.

(73) a. *Kue* *sõbe* *=d'e* *ze-i-pe?*
rain how.much =LOC come-IRR-Q
When will it rain? [lit. in how much will rain come?]

b. *Kue* *nu* *ze-ya.*
rain tomorrow come-FUT
It will rain tomorrow. [lit. rain will come tomorrow.]

(74) * *Kue* *nu* *=ta* *ze-ya.*
rain tomorrow =FOC come-FUT
It will rain tomorrow. [lit. rain will come tomorrow.]

In summary, in addition to the default focus position, Katío also marks focus with *ta*. The focus marker *ta* marks narrow focus on absolutive constituents but cannot mark narrow focus on ergative or oblique constituents.[2]

The fact that *ta* attaches only to absolutive arguments makes it appear to be an absolutive case marker, as Rex, Schöttelndreyer and Mortensen have all posited. Rex and Schöttelndreyer's analysis that it is a case marker is shown to be problematic by examples such as (61) and (67) in which marking an absolutive argument with *ta* resulted in an

[2] Here I note two possible occurrences in my data of *ta* on an oblique of time, which appear to contradict the claim that it only occurs with absolutives. However, these occurrences are not weighty compared to the many obliques of time in the data corpus that are not marked with *ta*. In example (i), it appears to mark *nruema* 'next day' or 'every day'. This example comes from a text published in the Ethnolinguistic Dictionary (Villalobos & Garcia 2004:58). However, the language consultant from Dabeiba, Antioquia did not recognize the word *nruemata* or the use of *ta* in this case. It is likely that *nruemata* is a single morpheme, as suggested by its entry in Villalobos and Garcia (2004:58) where it is listed as a variant of *nruema*.

(i) *Dai-rã* *nruemata* *do* *ed'a* *wã* *pan-a-ri-a.*
1PL.EXCL-PL every.day river LOC go AUX.PL-NPRS-HAB.PRS-DECL
We go to the river every day. [V&G:58]

In another isolated case, shown in example (ii), *ta* affixes to the phrase *ewari jõma* 'every day', and the language consultant who recorded the text indicated that it has a "stronger" meaning. This seems more likely to be an actual case of *ta* 'FOC' than the case in (i), but it is an isolated example. Further research will confirm or challenge the distribution of *ta* described in this study.

(ii) *Bu* *=a* *buru* *dai* *ewari* *jõma* *=ta* *bu* *bedea* *=ta* *dai* *=d'e* *ed'a*
2SG =ERG head 1PL.EXCL day all =FOC 2SG word =FOC 1PL.EXCL =LOC LOC

b'ʉ-tua *ewari* *jõma* *=ta.*
keep-IMP day all =FOC
You, in our heads, every day, keep your word in us every day. [Prayer:2]

ungrammatical utterance. Mortensen's analysis is more compelling, taking into account its pragmatic status function, yet his presentation suggests that he considers it to be a case marker fused with focus (1999:49).

The absolutive distribution of *ta* shown in this section does not automatically mean that it functions as an absolutive case marker. If *ta* marks focused information, then we can explain its mutual exclusivity with ergative constituents in discourse by referring to patterns of information flow. Cross-linguistically, new participants are rarely introduced in the position of agent of a transitive clause. Rather, they are frequently introduced in an existential clause, as a single argument of an intransitive clause, or patient of a transitive clause. DuBois, in a study of Sacapultec Maya, states that the "single new-argument mention typically appears in S or O [P] roles, but not in A" (1987:828). In narrative, by the time a participant has been established in the discourse and appears as the agent of a transitive clause taking ergative case, it has already been incorporated into the mental representation and is unlikely to appear as focused information. This causes ergative case and focus to only rarely coincide in cases such as the questions and answers above.

Velupillai suggests that the pragmatic motivation for an ergative system has to do with focus: "what ties S and P together... are the semantic notions of change of state and the pragmatic roles of FOCUS" [emphasis in original] (2012:253). Thus we may expect a focus marker to coincide closely with the absolutive case and to be practically mutually exclusive of the ergative case. This is a possible motivation for the fact that focus coding *ta* does not occur on ergative elements, although it is less helpful with explaining its absence on oblique elements.

It is worth noting that *ta* is not strictly required in the questions above with the focused absolutives marked with *ta*. The *ta* marking narrow focus on the patient argument *k'ãrẽ* 'what' in (45) and the S argument *k'ai* 'who' in (46) is optional in both cases, as indicated by the grammaticality of the following questions in (75) and (76). These are acceptable forms to my language consultant, but the forms marked with *ta* are preferred.

(75) *Marã* *k'ãrẽ* *k'o* *nu-m-u-pe?*
2PL what eat PROG-be-PRS-Q
What are you eating?

(76) *K'ai* *jũẽ-si* *puwuru* *ed'a* *=ra?*
who arrive-PST town LOC =TOP
Who arrived in the town?

The presence or absence of *ta* in these cases of focused absolutive arguments could be a matter of style and emphasis, depending on the speaker. My language consultant could not articulate any difference in meaning between a question with a focus marker and a question with no focus marker, although further studies of the verbal system and how it affects argument marking may be revealing. Since these question words are already in the default focus position, adding *ta* redundantly marks focus, yet marking focused absolutive words with *ta* is consistently preferred.

3.4 *ta* and complement clauses

In addition to marking focused words and phrases, the focus marker *ta* also marks focused complement clauses.

A complement clause may appear in the place of an absolutive constituent as an object complement. As in (77), it often appears after the verb rather than in the pre-verbal default focus position, and it is usually marked with *ta*. Example (77) shows a complement clause marked with *ta* as the patient of the main verb *ũdu* 'see'.

(77) *Mu* *=a* *nuweda* *ũdu-si-a* *[chĩna* *bea* *nu]* *=**ta**.*
1SG =ERG yesterday see-PST-DECL pig kill PROG =FOC
Yesterday I saw (you) killing a pig.

The verb *k'awua* 'know' may also take a complement clause as an argument. In (78), the object complement occurs again post-verbally and is marked with *ta*.

(78) *Aramãũne* *k'awua-si-a* *[wãrĩnu* *izhi* *DaiZeze* *=ba* *zok'a-da]* *=**ta**.*
just.then know-PST-DECL true 3SG God =ERG send-COMPL =FOC
So then [I] knew that truly God had sent it. [BBT:125]

Although it is most common for the complement clause to follow the verb and be marked with *ta*, it may also precede the verb, in which case it may be unmarked because it in the default focus position, as in (79).

(79) *...[k'ai =ta aʉdre zarea chu b'-ʉ] k'awua-d'a =ita...*
who =FOC more strong STV AUX-PRS know-PL =PURP
...So that [we] will know who is stronger... [V&G:p131]

If a complement clause is not in focus, it is not marked with *ta*. Given the context of the question in (80a), example (80b) is ungrammatical because the complement clause *kue zei* is not in focus. (80c) shows the grammatical response to (80a), where *ĩjã ẽã* 'don't believe' is the focused portion of the utterance.

(80) a. *Kue ze-i-ka?*
rain come-IRR-Q
Will rain come?

b. * *[Kue ze-i] =ta ĩjã ẽ-ã.*
rain come-IRR =FOC believe NEG-DECL
[I] don't believe rain will come.

c. *Ĩjã ẽ-ã [kue ze-i] =ra.*
believe NEG-DECL rain come-IRR =TOP
[I] don't believe rain will come.

Indirect quotations like example (81) usually occur with the verb *jara* 'say'. They are treated like absolutive complement clauses, and are often marked with *ta*. In (81), since the complement clause is equative and there is no verb, *ta* attaches to the rightmost constituent, *papa* 'mother'.

(81) *Zõrã-rã =ba jara-si-d'a [jedeko =ra kue papa] =**ta**.*
old.person-PL =ERG say-PST-PL moon =TOP rain mother =FOC
The old people said that the moon is the rain's mother. [V&G:p87]

Direct quotations usually occur with the verb *a* 'say', which follows the quotation as in (82), or with *mawũã* 'say thus' which precedes it, or with both. Direct quotations are not marked with *ta*.

(82) *Aramaũ̈ta mɯ̈ =a "mawũẽ mɯ̈ =a =ra kewara aku-ze-ya"*
just.then 1SG =ERG then 1SG =ERG =TOP afternoon look-come-FUT

a-si-a.
say-PST-DECL
Then I said, "Then I'll come see in the afternoon". [Africanas:7]

In all the previous examples of complement clauses that contain a verb, the verb is clause final. However, when the complement clause contains a verb that is not clause final, *ta* must attach to the verb instead of to the rightmost element, as illustrated by the complement clause of the verb *k'awua* 'know' (83) below.

(83) *Aramaũ̈ne mɯ̈ =a k'awua =ita [[mɯ̈ izhi bedea jara wã-ĩ bara]*
just.then 1SG =ERG know =PURP 1SG 3SG word say go-IRR OBLG

*=**ta** krĩña b'-ɯ̈ =**ta** izhi =a].*
=FOC want AUX-PRS =FOC 3SG =ERG
So that way I would know if He wanted me to go preach. [BBT:33]

Example (83) is notable in that it contains two complement clauses marked with *ta*. The first complement clause, marked by the outermost brackets, is the object of the verb *k'awua* 'know': 'So that way I would know [if he wanted me to go preach]'. The second complement clause, marked by the innermost brackets, is inside of the first. It occupies the patient slot of the first complement clause and the verb *krĩña* 'want': 'he wanted [me to go preach]'

In the first complement clause, the constituent *izhi=a* "3SG-ERG" has been postposed, and thus is the rightmost element, farther right than the verb which hosts the focus marker. Since *ta* cannot attach to an ergative case argument, it is pushed leftward onto the verb. The result is a complement clause marker that is clause-internal. Or, alternatively, if we consider a postposed argument to be core-external, then we may state that *ta* attaches to the final constituent of core.[3]

Some further functions of *ta* and other focus morphemes are described in the following section.

[3] The notion of core is explained in Van Valin and LaPolla (1997).

3.5 Unmarked and marked focus

In a simple SOV transitive clause such as (84), the O is in the default focus position and typically (but optionally) is marked with *ta* even when there is no apparent contextual motivation for narrow focus.

(84) *Wũãwũã =ba be =ta jida-si-a.*
child =ERG maize =FOC grab-PST-DECL
The child grabbed the maize.

The sentence in (84) is an isolated elicited sentence and is not answering the question "what did the child grab?" which would put the constituent *be* 'maize' in narrow focus. Yet even though this could be a predicate-focus or even a sentence-focus utterance, and even though there is no apparent contextual motivation for *be* 'maize' to be in focus, it is still marked with *ta*.

I conclude that the focus indicated by *ta* is relatively neutral, not unlike the unmarked narrow focus of the default focus position. Marked narrow focus is indicated with different morphemes.

Just as the narrow focus morpheme *ta* attaches only to absolutive elements and not to non-absolutive elements, there is also a dichotomy in marked narrow focus between the treatment of absolutive and non-absolutive elements.

For this analysis, I have reframed Mortensen (1999:146), who puts the use of *ta* in terms of activation status, indicating that *ta* marks focus on a new participant, with "introductory focus." According to him, an absolutive that is already activated or "given" in the discourse is focused with *tru*, and an activated or "given" "nonabsolutive" constituent with *buru*.

I frame the analysis not in terms of activation status, but of marked and unmarked focus (Van Valin & LaPolla 1997:209). The focus marker *ta* shows unmarked focus, parallel to the function of the default focus position, and *tru* and *buru* show marked focus, what is often termed "emphasis," as in Rex (1975:40). Marked narrow focus on an absolutive element is marked with *tru*, and marked narrow focus on a non-absolutive element is marked with *buru*, both glossed below as 'EMPH'.

Example (85) shows absolutive case *maũ* 'that'[4] receiving marked narrow focus *tru* in NEP.

(85) *Kĩru-buru =ba jara-si-a "maũ =**tru** mu =a ũrĩ*
get.angry-PRS =ERG say-PST-DECL that =EMPH 1SG =ERG hear

kĩrĩa-ka".
want-NEG.HAB
Out of anger he said, "I don't want to hear ***that***!". [Mortensen 1999:146]

Example (86) shows *buru* marking ergative element *Dazhi Zhibari* 'God' and in example (87) it marks *naũ* 'this'.

(86) *Mu =a k'awua ẽ baera, Dazhi Zhibari =**buru** k'awua b'-u-a...*
1SG =ERG know NEG because God =EMPH know AUX-PRS-DECL
Because I don't know, ***God*** knows... [Gossip:10.2]

(87) *Chidima =ra mesẽra wã ze toto-si, wabema-rã =ba jara-si-d'a:*
grey.squirrel =TOP quickly go come go-PST other-PL =ERG say-PST-PL

*"Naũ =**buru** dazhi-rã =á bania ũdu-bi-ya".*
this =EMPH 1PL.INCL-PL =DAT water see-CAUS-FUT
The grey squirrel comes and goes quickly. The others said: "***This one*** will show us the water." [V&G:p90]

In the case of ergative noun phrases such as in examples (86) and (87) above, *ba* is omitted altogether and only *buru* appears. On the other hand, when *buru* is affixed to an ergative pronoun that takes the ergative marker allomorph *a*, it is added to the ergative marker as in *mu=a=buru* '1SG=ERG=EMPH' of (88) (Mortensen 1999:48). Each clause in (88) is marked with *ta* because they are complement clauses of the verb *krĩcha* 'think' in this example where the speaker describes what others wrongly thought of him.

[4] The demonstrative *maũ* may be translated as 'this' or as 'that' in different examples in this thesis. Instead of referring to physical proximity or distance, like *naũ* or *kaũ*, respectively, *maũ* refers to entities that are out of sight of the speaker or not physically present (Harms 1994:45).

(88) *...mʉ* **=*a*** **=*bʉrʉ*** *merã* *diabue* **=*ta*,** *mʉ* =*a* =*bʉrʉ* *plata* *dia-si*
1SG =ERG =EMPH hide send =FOC 1SG =ERG =EMPH money give-PST

=*ta*...
=FOC
...that ***I'm*** sending [her] in hiding, that ***I*** gave [her] money... [Gossip:6.1]

In summary, Embera Katío has two strategies for showing unmarked narrow focus: the pre-verbal default focus position and the focus marker *ta*. For absolutive elements, these two strategies typically co-occur, with the constituent in the pre-verbal position also being marked with *ta*. Although *ta* is not permitted to mark unmarked narrow focus on ergative or oblique elements, their placement in the pre-verbal default focus position achieves this. Stronger highlighting is achieved with the marked narrow focus morphemes *trʉ* for absolutive and *bʉrʉ* for ergative.

CHAPTER 4

TOPIC MARKER *RA*

I begin the discussion of the functions of *ra* by reviewing the analyses offered by other researchers in 4.1. Then I discuss the definition and usage of the term topic in 4.2, the analysis of two *ra* morphemes versus one in section 4.3, and finally show examples of the distribution of *ra* and how it can be described with the concept of topic. I conclude that *ra* is best analyzed as a single morpheme: a topic marker that marks elements that pertain to the pragmatic topic, including syntactic topics and postposed elements.

4.1 Previous analyses of *ra*

As briefly introduced in section 1.2, the function of *ra* in the Embera languages has been described in many different ways. Loewen calls it an "emphatic enclitic", undistinguished in function from *ta* and *trɨ* (1958:101).

Rex (1975:39-40) refers to *ra* as a "non-case marker" of previous reference: "*-ra* indicates that the participant being referred to is one previously mentioned... It may combine with ergative *-ba* forming *-bara*, which has the same distribution as *-ba*."

Schöttelndreyer (1977:152) describes *ra* as marking the end of a conflict in the narrative. She writes that the presence of *ra* "signals... that the conflict is over" or that the participant marked with *ra* is "about to perform actions pertaining to the conflict" (1976:256). However, she also cites Rex's analysis in an endnote in the Spanish translation.

Aguirre Licht (1999:326-327) called the Embera Chamí (SE) *-ra* a topic marker suffix, whose function is to "topicalize" a "nominal" entity, that is, "a stable entity, a place, a time, a nominalization or, in general, what can be understood as a state of 'stative things or situations'". According to him, an element marked with *ra* may occupy any position in

the sentence. He does not define topicalization, however, except to say that it is "another procedure for expressing the promotion or highlighting of an element of the sentence."

Harms also describes a *ra* morpheme in Epena Pedee (SE) which "indicates a switch to the entity that is focused on... from a previously stated or implied entity to which it corresponds..." (1994:82). According to him, it indicates "dissimilar comparisons, intensified plurals, switches of focus and conditionals." In an example in which it occurs with the ergative *pa*, he describes *ra* as "comparative" and marking a "shift in attention" (1994:190).

In Embera Tadó (SE), the function of *ra* is to mark a shift in referent, labeled 'pivot', according to linguist David Pickens and native speaker Carlos Duave (personal communication).

In Mortensen (1999), two different suffixes with the form *ra* are described. The first morpheme is a "normal" or "nonfocal" absolutive case marker that attaches to absolutive arguments (1999:143). Example (89) below illustrates a sentence in which *ra* on an absolutive argument *mʉ* '1SG' corresponds to Mortensen's first "absolutive case marker" *ra* morpheme.

(89) *Aramãũne mʉ =**ra** puwʉrʉ Uré a-ba-da ed'a wã-si-a.*
just.then 1SG =TOP town Uré say-HAB-PL LOC go-PST-DECL
So then I went to a village called Uré. [BBT:72]

His description of the second *ra* in Embera Katío is that it is a "pivot" suffix that attaches to an ergative noun phrase with *ba* or to other non-absolutive elements in a clause, indicating a shift in attention away from the constituent it marks. According to Mortensen, this *ra* suffix is used to "direct attention away from the constituent to which it is attached to the next corresponding constituent, so may be considered to be a marker of anticipatory or cataphoric focus" (1999:145). He further explains that "*-ba-ra* marks a polar construction: what follows a noun suffixed with *-ba-ra* is emphasized." It is this morpheme that occurs on elements other than nouns, "such as dependent verb forms, locatives, demonstrative adverbs, etc." (personal communication).[1]

[1] Through personal communication with Mortensen, I understand that in contrast to when his grammar was published, he now believes the two *-ra* forms to have the same function of shifting focus away from the constituent on which it occurs. Thus he considers them to be the same *ra* morpheme, whose most frequent function in NEP is to mark absolutive case.

As described by Rex and Mortensen, in Embera Katío, *ra* is very frequently attached to an ergative noun phrase with *ba*, for a combined noun phrase marking of *bara*, as on *Daizeze* 'God' in example (90). This usage of *ra* with an ergative argument corresponds to Mortensen's second "pivot" suffix. In example (90), I have glossed *ra* as 'TOP' in line with my own analysis, but Mortensen's text shows *ra* here glossed as 'PIVOT'.

(90) *DaiZeze* =*ba* =***ra*** *awara* *izhi* =*a* *krĩna* =*ta* *o-marẽã* *zok'a-si*
God =ERG =TOP rather 3SG =ERG want =FOC do-PURP send-PST

baera *mawũã* *mĩõ* =*ba* *mɨ̃ɨ̃* =*á* *plata* *dia-d'a* *ẽ*
because like.this nobody =ERG 1SG =DAT money give-PL NEG

b'-a-si-a.
AUX-NPRS-PST-DECL
Because God had sent me to do His will, no one gave me any money at all. [BBT:76]

In summary, there are three main groups of terminology for *ra*: "previous reference" from Rex; "pivot/shift in attention", from Harms, Pickens and Mortensen; and "topic" from Aguirre Licht.

As seen above, in addition to these three groups, Mortensen describes another *ra* morpheme as a nonfocal absolutive case marker, which implies two different *ra* morphemes. Calling one *ra* an absolutive case marker automatically requires that there be two different morphemes, because by definition an absolutive case marker cannot occur on an ergative argument like it does in (90). However, I do not believe it is necessary to posit two different morphemes in order to adequately describe the function of *ra*, and it is preferable to posit one meaning for one form wherever reasonable (Reesink 1981:91). All that is necessary to posit a single *ra* morpheme is to remove its classification as an absolutive case marker. Then there is no trouble accounting for its frequent occurrence on ergative and oblique arguments. In section 4.3, I discuss why the notion of topic adequately accounts for the occurrences of *ra* in Embera Katío that Mortensen (1999) describes as absolutive case markers.[2]

[2] Note that my conclusions apply only to Embera Katío, and the use of *ra* may differ in NEP. Mortensen indicates that the use of *ra* in NEP is distinct from that of Embera Katío, perhaps more distinct than Mortensen (1999) implies (personal communication).

I suggest that the terms previous reference and pivot are referring to the same notion. Previous reference marking is meant to identify an argument as being old information and not requiring the focus and attention due to a new argument. In other words, it directs the focus and attention elsewhere in the sentence, which is reminiscent of the definitions of pivot given by Harms, indicating "a switch to the...focused [entity] from a previously stated... entity...." Harms (1994:82) and Mortensen, to "direct attention away from the constituent to which it is attached to the next corresponding constituent..." (Mortensen 1999:145).

This combined function of marking a previously identified argument and directing attention to another part of the utterance is a property of the pragmatic topic. In the following section, I define the notion of topic as used in this thesis.

4.2 Notion of topic

As Payne (1997:270) notes, the term topic has been used in numerous ways by researchers. He even notes that "the terms 'focus' and 'topic' are antonyms in some traditions and synonyms in others" (1997:262). For this study, I use the term topic on two distinct levels: the pragmatic topic and the syntactic topic (following Gundel (1988:210-211)), each of which I discuss in more detail below. It does not refer to the overall "theme" of a discourse, as we may use it in casual speech (e.g. "What was the topic of his speech?").

The pragmatic topic is defined by Gundel as follows: "An entity, E, is the topic of a sentence, S, iff in using S the speaker intends to increase the addressee's knowledge about, request information about, or otherwise get the addressee to act with respect to E." She also describes it as "the domain within which the main predication holds" (1988:210). The pragmatic topic is the part of the utterance that is the familiar or identifiable base or domain to which the speaker adds the focused comment.

The pragmatic topic closely correlates with, and therefore is easily confused with, the referential properties of discourse activated vs. nonactivated and given vs. new information. The topic is most likely to be an entity that is already active in the discourse or is at least accessible, and is unlikely to be a brand-new participant unconnected to any other established referent. Van Valin and LaPolla present a topic acceptability scale where an

active referent is the most acceptable topic, followed respectively by an accessible referent, an inactive referent, a brand-new anchored referent, and finally the least acceptable topic is the brand-new unanchored referent (1997:204). A brand-new referent is much more likely to be a focus and is very unlikely to appear as a topic.

In the same way, the topic of a topic-comment sentence *prototypically* refers to old or given information (Velupillai 2012:233). Example (91) comes from the middle of a text describing the attributes of frogs, and so the referent of the pronominal syntactic topic *izhi* '3SG' is given from the discourse context, and so in that sense, the referent is "activated" and "given". This element specifies the domain of application of the focus: the fact that there are no teeth. This is an example of a topic that corresponds to "activated" and "given" information.

(91) *Izhi* =*ra* *kida* *ne* *ẽ-ã.*
3SG =TOP tooth thing NEG-DECL
As for it [the frog], there are no teeth. [V&G:p41]

Thus there is a high correlation between topics and referents that have been previously mentioned, which accounts for Rex's analysis of *ra* as a marker of "previous mention." However, in spite of the high correlation between the pragmatic topic and activated or old information, it is not a perfect correspondence. Gundel defines the topic with conditions of familiarity and identifiability, which I find more useful, but notes that these too have occasional exceptions (1988:215).

For the purposes of this thesis, I define a "topical" element as being related to or forming part of the pragmatic topic. In the discussion of *ra*, I refer to elements being marked as topical, meaning that they form part of the pragmatic topic, and are not in focus. This will in many cases correspond with the notions of "previous mention," "activation," or "given," but should not be understood as synonymous with these.

The second use of the term "topic" is the syntactic topic, referring to the concept of the topic-comment sentence form. Topic-comment sentence articulation stands in contrast to subject-predicate articulation (Li & Thompson 1976), and refers to a two-part sentence structure: first, the topic defines the domain of the comment's assertion and second, the

comment is what is asserted about that domain. Embera Katío has sentences with both subject-predicate articulation and topic-comment articulation.

Examples (91), (92) and (93) are examples of topic-comment articulation in Katío. The topic of these structures is consistently marked with *ra*.

(92) *Adichichi* =*ra* *kĩrã̃* *bio* *bara.*
cricket =TOP face many be.many
As for cricket(s), there are many kinds. [lit. As for cricket(s), there are many faces.]
[V&G:p10]

(93) *Dai* *drua* *trʉ̃* =*ra* *Mobʉ* *a-ba-d'a.*
1PL.EXCL land name =TOP Mobʉ say-HAB-PL
Our land's name is called Mʉra [lit. Our land's name, they call [it] Mʉra.] [V&G:p10]

Dooley and Levinsohn note that topics "In general... will be definite or generic, rather than indefinite" (2001:69). This corresponds to Gundel's condition of identifiability: an indefinite referent is not identifiable (1988:214). In example (92), the topic is generic, and in example (93), the topic is identifiable.

Thus in a prototypical topic-comment sentence, the two notions of topic overlap: the syntactic topic is always pragmatically topical. "While syntactic topics always refer to pragmatic topics, a pragmatic topic is not always encoded as a syntactic topic" (Gundel 1988:211). In keeping with this, Embera Katío marks syntactic topics with *ra*, and also marks other elements that are not in the syntactic topic position as topical. Evidence for this is shown in sections 4.4.1 and 4.4.2.

To illustrate the notion of topic at work in a language, I refer to Japanese. The use of the Japanese postpositional markers *wa* and *ga* depends on the notion of topic. In a topic-comment structure which has predicate focus, and in which the topic is assumed to be familiar and identifiable to the listener based on the discourse, the topic will be marked with *wa* as in (94a). However, in a sentence-focus utterance where none of the information is assumed to be familiar and identifiable (i.e. there is no topic), *ga* will be used instead. I have underlined the focus in example (94), which is from Lambrecht (1994:223):

(94) a. *(Kuruma wa)* *kosyoo-si-ta*
(car TOP) break.down-do-PST
(The car), it broke down (in response to "What happened to your car?") - Predicate focus

b. *Kuruma* *ga* *kosyoo-si-ta*
car SBJ break.down-do-PST
My car broke down (in response to "What happened?") - Sentence focus

Because of this distinction, Japanese *wa* is often referred to as a topic marker. In these examples, the use of *wa* indicates a topic-comment structure, while *ga* indicates a subject-predicate structure. Both sentence articulations are prominent in Japanese (Li & Thompson 1976:460).

Payne refers to Japanese as an "overlay" system in which *wa* can replace the subject marker (95b), object marker (95c) or other case markers when a certain constituent is "singled out for special pragmatic treatment" and made into the syntactic topic. Example (95) is from Payne (1997:278):

(95) a. *Taroo ga hon o katta*
Taro SBJ book OBJ bought
Taro bought a book

b. *Taroo wa hon o katta*
Taro TOP book OBJ bought
As for Taro, he bought a book

c. *Hon wa taroo ga katta*
book TOP Taro SBJ bought
As for the book, Taro bought it

Like the Japanese topic marker *wa*, Embera Katío topic marker *ra* can also occur on noun phrases in various cases. However, in contrast to Japanese, the topic marker is not "overlayed" by replacing the case marker, but instead occurs to the right of the case marker, as in example (96), where it occurs with an ergative case-marked generic subject, and (97), where it occurs with a dative case-marked definite noun phrase.

(96) *Bok'orro* =***ba*** =***ra*** *k'o-ba-ri-a* *egoro* =*ta...*
frog =ERG =TOP eat-NPRS-HAB.PRS-DECL earth =FOC
The frog eats dirt. [V&G:p42]

(97) *DaiZeze* =*ba* *k'ãrẽ* =*ta* *jara-si* *dama* *torro* =***á*** =***ra?***
God =ERG what =FOC say-PST snake white =DAT =TOP
What did God say to the white snake?

The free translations of the examples in (95) above highlight the common strategy for translating Japanese noun phrases marked with the topic marker *wa* into English: "As for Taro..." or "As for the book...." This structure is the closest approximation in English to topic-comment sentence articulation and a morphologically-marked topic. For example: "As for my cat, she's getting better." For this topic construction in English, the element "my cat" is left-dislocated, marked with "as for," and is followed by the remark that contains the focus: "she's getting better."

This device is mainly (but not frequently) used in English to establish or change a topic before adding new information, and it is not necessarily the case that the cat in the example above has already been introduced into the discourse. It is more likely that "my cat" is being introduced in contrast to another previously mentioned referent. Chafe notes that this construction often implies contrastiveness in English (1976:49). For this reason, an English free translation that uses this device to approximate topic-comment structure should not be considered exactly equivalent to what is indicated by *ra* in a given example. However, with understood limitations, this topical translation "as for..." may be useful to translate and better understand some instances of *ra* in Embera Katío. Example (98) is one such instance, where *ra* marks the initial topic of the whole utterance, *bok'orro* 'frog', and then marks the topic *ũkuru* 'some' of each successive clause.

(98) a. *Bok'orro* =*ra* *ũkuru* =*ra* *arabia* *chu* *b'-u-a,*
frog =TOP some =TOP bigger STV AUX-PRS-DECL
As for frogs, some are bigger,

b. *ũkuru* =*ra* *wĩka* *kir-u-a,*
some =TOP little be.DIM-PRS-DECL
some are small,

c. *ũkuru* =*ra* *kuara* *chu* *b'-u-a.*
SOME =TOP yellow STV AUX-PRS-DECL
some are yellow. [V&G:p42]

Another language that offers an example of topic morphology is Huallaga Quechua of Peru (Weber 1989). Its suffix *-qa* has a wide range of functions, one of which is to mark information that is previously mentioned or accessible from information already mentioned. It is not used to introduce new information. Weber calls it topic for consistency with the tradition within Quechua studies, but emphasizes the arbitrary nature of that label, suggesting that "thematic" might be a more appropriate term. It fits the general description of the pragmatic topic: "*-qa* tends to occur on constituents that have been previously mentioned or alluded to, or are part of the "general knowledge"" (1989:400). It marks the connection between the new utterance and its pragmatic context: "*-qa* occurs on the constituent(s) most responsible for a sentence's relevance to its context, for it is the function of these words to express the logical or temporal relationship of the sentence to what precedes it in the discourse" (1989:412). I suggest that *ra* has a pragmatic function similar to that of *-qa* of Huallaga Quechua described above.

Weber proposes a basic function of *-qa* which he clarifies as just a "step toward an adequate characterization:" "*-qa* occurs on those constituents of a sentence which the speaker wishes to indicate as most responsible for the sentence's relevance to its context" (1989:404). A similarly basic description, just as open to further revision as Weber's, could apply to the function of *ra*.

Weber says of *-qa*, "There is little or no restriction on the occurrence of *-qa* in terms of the syntactic category to which it is suffixed; it occurs with elements of all the major syntactic categories," but not with finite verbs except in highly restricted and rare cases. However it is limited to the "main constituents" of a sentence, e.g. it cannot occur inside of a subordinate clause, or a noun phrase, etc. (1989:393-395).

The morpheme *ra* in Katío behaves similarly, with little to no restriction for the syntactic categories with which it occurs (pronouns, nouns, adjectives, adverbs, verbs, postpositions, numbers, negatives, etc.). However, it always occurs to the right of a phrasal unit, either a noun phrase, a postpositional phrase, or a subordinate clause of condition or time, as shown in section 4.4. However, in contrast to *-qa*, *ra* can occur inside a complement clause as in (99), repeated from (81), where it marks the syntactic topic of the complement clause.

(99) *Zõrã-rã* =*ba* *jara-si-d'a* *[jedeko* =***ra*** *kue* *papa]* =*ta.*
old.person-PL =ERG say-PST-PL moon =TOP rain mother =FOC
The old people said that the moon is the rain's mother. [V&G:p87]

The distribution of *-qa* and *ra* is consistent with Gundel's description of topic markers: ""Topic markers are typically not restricted to marking any one syntactic relation or semantic role... they can mark any nominal constituent, including one that is not an argument of the verb at all..." (1988:216).[3]

To summarize this discussion as it applies to *ra*, I suggest that *ra* marks elements of the pragmatic topic. This means that it also marks the syntactic topic, since Gundel states that "syntactic topics always refer to pragmatic topics" (1988:211). Thus when I refer to *ra* as a topic marker, that term includes the notions of pragmatic topic, and by implication, syntactic topic.

4.3 Absolutive case and *ra*

In this section, I discuss why positing two *ra* morphemes is unnecessary since we can account for all cases of *ra* with the notion of topic.

The topic marker *ra* frequently occurs on absolutives, but if it were a "nonfocal" absolutive case marker as proposed by Mortensen, we would expect it to occur regularly on any activated absolutive case constituent that is not in focus. In the following exchange between the King and an Embera man, *k'abayo* 'horse' is introduced in the discourse in (100a) where it, together with its relative clause 'that I have', is marked with focus marker *ta*. It is mentioned again in (100c), where it is again the absolutive object of the transitive verb *zr~~u~~ga* 'steal', but is unmarked. Here, *k'abayo* is previously activated in the discourse, so its lack of marking cannot be explained as a nonactivated -ø absolutive, as in Mortensen (1999:49). *K'abayo* is not in focus either, since it is clearly "predictable" and "recoverable," from the discourse since in (100b) it was omitted altogether.[4] As an activated, nonfocal

[3] Gundel's description of a topic marker marking "any nominal constituent" is perhaps sufficient motivation to analyze the adverbial obliques and postpositional phrases that may be marked with *ra* as nominal. I have found no evidence in my data to motivate a syntactic class of adverb, and more evidence would be needed to motivate a syntactic class of postposition.

[4] Recall Lambrecht's definition of focus: "The focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the UNPREDICTABLE or pragmatically NON-RECOVERABLE element in an utterance" [emphasis in original] (1994:207).

element, this is where we would expect a "nonfocal absolutive marker" *ra* to appear, but instead, the absolutive argument in (100c) is unmarked.

(100) a. King: *Mobe =ra, [k'abayo mɯzhi =a animi] =ta zrɯga-rua!*
then =if horse 1SG =ERG have =FOC steal-IMP
If that's so, steal [the horse that I have]!

b. Embera: *Me, mɯ =a zrɯga-ya.*
alright 1SG =ERG steal-FUT
Alright, I'll steal [it]

c. Embera: *K'abayo zrɯga-ya.*
horse steal-FUT
I'll steal the horse.

d. Embera: *Mobe =ra, nu, nu zrɯga-ya mɯ =a k'abayo*
then =if tomorrow tomorrow steal-FUT 1SG =ERG horse

=ra.
=TOP
If that's so, tomorrow, tomorrow I'll steal the horse. [EP:104-108]

If *ra* is a pragmatic status marker of topicality that marks elements that pertain to the pragmatic topic (thereby signaling that some other element in the sentence is the focus), then we would expect it to occur in certain topical constructions and optionally on other constituents according to the information flow and perhaps according to the stylistic preference of the speaker. This is what we find in (100), where *ra* fails to appear on the activated nonfocal absolutive argument in (100c), but does appear in (100d) on the same argument.

In (100d), *k'abayo* 'horse' occurs in a postposed construction, while the oblique *nu* 'tomorrow' is in focus, being repeated and occurring in the pre-verbal focus position. This is the sort of situation in which we would expect a marker of topicality to occur: in a construction which contributes to the focal highlighting of another constituent. Here, marking the post-verbal constituent as topical with *ra* is redundantly marking it as not in focus, much like pre-verbal focused absolutives are redundantly marked with *ta*.

Finally, examples like (101) below are difficult to account for if *ra* is an absolutive case marker, since in (101b) we have two noun phrases: one with *ra*, *kima* 'spouse', and one unmarked, *trɯ̃* 'name'.

(101) a. *Mʉ zeze trʉ̃ Zozoareame.*
1SG father name Zozoareame
My father's name is Zozareame.

b. *Zhi kima =ra trʉ̃ Cildaba-a*
3SG spouse =TOP name Cildaba-DECL
As for his wife, [her] name is Cildaba

Which is to be understood as the absolutive subject of the sentence? A reading of *zhi kimara trʉ̃* as a single possessive noun phrase "his wife's name," like the phrase *mʉ zeze trʉ̃* in (101a), is not possible because it is divided by *ra* occurring after *kima*, marking the rightmost edge of the noun phrase. Gundel states that topic markers "are always postpositional and typically follow a sentence-initial constituent," as when *ra* marks a syntactic topic. "They thus serve to mark the major constituent boundary between topic and comment" (1988:222).

A correct interpretation is only possible when we see this in terms of a topic-comment construction, with *ra* marking the boundary between the topic and the comment. This establishes *zhi kima* as the topic, accessible based on the previous introduction of the speaker's father. The comment is then made up of the equative clause "[her] name is Cildaba", whose absolutive subject is *trʉ̃* 'name'.

Calling *ra* a topic marker allows us to unite all its occurrences under a single notion without positing a connection between *ra* and the absolutive case. Section 4.4 shows examples of the distribution of *ra* and discusses how its function as a topic marker accounts for its distribution in Embera Katío.

4.4 Distribution and functions of *ra*

In this section, I discuss *ra* marking the syntactic topic in 4.4.1, the pragmatic function in 4.4.2, and finally how postposed elements are usually coded as topical, marked with *ra* because of their relation to the pragmatic topic.

4.4.1 Syntactic topic

First I present examples of *ra* marking the syntactic topic, which may be a noun phrase, a postpositional phrase, or a subordinate clause of condition or time.

Example (102) is made up of sentences 1, 2 and 8 from a text about the river, which is central to Katío culture. In (102), *ra* marks two elements: the syntactic topic, *do* 'river', and a benefactive element *mɯ̃=ita* 'for me'. Both of these serve as topical information to which the following comment "is like a sibling" applies. Since *do* 'river' is both the topic and subject of the comment, the topic-comment structure of the sentence is not entirely clear. The sentence could also be interpreted as a subject-predicate structure such as "For me the river is like a sibling." However, *do* 'river' in (102c) is a clear example of a syntactic topic, followed by the comment "some are angry."

(102) a. *Mɯ̃ =a mejacha kɯ̃ãg'a b'-ɯ̃-a do =**ra**.*
1SG =ERG much love AUX-PRS-DECL river =TOP
I love the river a lot.

b. *Do =**ra** mɯ̃ =ita =**ra** mɯ̃ mebea kĩrãk'a b'-ɯ̃-a.*
river =TOP 1SG =BEN =TOP 1SG sibling like AUX-PRS-DECL
The river, for me [it] is like a sibling.

c. *Do =**ra** ɯ̃kɯrɯ̃ =**ra** mesia chu b'-ɯ̃-a, mawasiudu dai-rã*
river =TOP some =TOP angry STV AUX-PRS-DECL even.so 1PL.EXCL-PL

pera-d'a ẽ-ã.
fear-PL NEG-DECL
As for river(s), some are angry, [but] even so we are not afraid. [V&G:p58]

The following example is from the text *K'õe k'õe* about the poisonous frog. It introduces its protagonist in (103a), and then describes more of its characteristics.

(103) a. *K'õe k'õe oi =d'e bema.*
poisonous.frog forest =LOC ORIG
The poisonous frog is from the forest.

b. *Izhi =**ra** dabu wĩka boremea kir-ɯ̃-a.*
3SG =TOP eye little protruding be.DIM-PRS-DECL
As for it, the eye is little and protruding.

c. *Izhi k'akua* =***ra*** *kuara chu b'-~~u~~-a.*
3SG body =TOP yellow STV AUX-PRS-DECL
Its body is yellow. [V&G:p128]

Example (103b) shows *ra* marking the syntactic topic *izhi* '3SG', followed by a comment that has its own subject, *dabu* 'eye'. In this case, *izhi* cannot be understood as the syntactic subject because of the presence of *dabu*. This is an example of a "double subject" construction typical of languages that favor topic-comment structure (Li & Thompson 1976:468).

By contrast, in example (103c), the element *izhi k'akua* 'its body' could potentially be understood as a syntactic topic or a syntactic subject that is marked as topical.[5]

The topic marker *ra* also frequently marks dependent clauses of condition. Gundel notes that it is typologically "common for topic markers to mark conditional and relative clauses, which share with topic the property of being given (presupposed) in relation to the rest of the sentence" (1988:218). Haiman states: "Conditionals, like topics, are givens which constitute the frame of reference with respect to which the main clause is either true (if a proposition), or felicitous (if not)" (1978:564).

When it occurs with the irrealis suffix *-i*, *ra* subordinates a clause with a conditional sense of 'if', such as in (104).[6] First, *b'ed'a* 'fish' is marked as the syntactic topic (it is also previously mentioned in the text, which is about how to catch fish). Second, the subordinate clause "they flee under the rock" is marked with *ra* as conditional. The information in this clause is topical in that, like a topic, it establishes a domain or "presupposition" for the focus.

(104) *B'ed'a* =***ra*** *[mog'ara edre miru toto-i]* =***ra*** *bi-ia bea*
fish =TOP rock under flee go-IRR =TOP very.good kill

pan-a-ri-a.
AUX.PL-NPRS-HAB.PRS-DECL
The fish, if they flee under the rock(s), are easily killed. [V&G:p33]

[5] Further research is needed to explore whether the presence of *ra* alone indicates that this is a topic-comment structure, or whether *izhi k'akua* could be interpreted as a syntactic subject marked as topical. Since a subject has a selectional relationship with the verb and a topic does not (Li & Thompson 1976:462-463), evidence that an element like *izhi k'akua* is subject to selection by the verb would point to it being a subject.

[6] This is not the only way to mark a conditional clause. Verbal suffix *-b~~u~~ru* is also a conditional subordinator.

In a similar case in example (105), *ra* marks a negative subordinate clause of condition 'if there were no river'.

(105) *Dazhi-rã [do ne ẽ-ã] =ra naʉ̃ drua =d'e b'-a b'e*
1PL.INCL-PL river thing NEG-DECL =TOP this land =LOC be-NPRS be.able

ẽ-ã.
NEG-DECL
We, if there were no river, wouldn't be able to live in this land. [V&G:p58]

Cross-linguistically, topic markers frequently mark time expressions (Gundel 1988:216). In Katío, the topic marker *ra* also subordinates adverbial time clauses that establish the time setting for the following comment. Example (106) shows an adverbial time clause *wũẽrãra ãbu zroma chubʉ =ra* "when a woman is big in the womb" and a condition clause with irrealis suffix *-i, euma ũdui =ra* "if [she] sees a rainbow". This example refers to the cultural superstition that pointing at a rainbow brings bad fortune.

(106) *Mʉ pazhõdra jara-si [wũẽrã =ra ãbʉ zroma chu b'-ʉ] =ra,*
1SG grandmother say-PST woman =TOP womb big STV AUX-PRS =TOP

[euma ũdu-i] =ra jara-dia b'e ẽ.
rainbow see-IRR =TOP say-give be.able NEG
My grandmother said that when a woman is big in the womb, if [she] sees a rainbow, she can't tell anyone. [V&G:p73]

The topic marker *ra* frequently occurs with locative *d'e* on a dependent time clause. In example (107), *ra* marks the dependent time clause 'when our stomach hurts'.

(107) *[Dazhi b'i pua kubʉ] =d'e =ra dazhi papa =ba*
1PL.INCL stomach hurt put =LOC =TOP 1PL.INCL mother =ERG

k'orrata =ta wẽ b'-a-ri-a.
lemon.verbena =FOC scrape AUX-NPRS-HAB.PRS-DECL
When our stomach hurts, our mother scrapes lemon verbena. [V&G:p29]

The examples in this section have shown *ra* marking the syntactic topic of a topic-comment sentence and subordinating clauses of condition or time which, like topic, are presupposed in relation to the following comment.

4.4.2 Pragmatic topic

I turn now to occurrences of *ra* that are not syntactic topics, but mark pragmatic topicality. As discussed in section 3.3, an ergative argument is unlikely to be in focus as new information and is likely to be a topical entity with some or all of the properties of familiar, definite, activated or given. Therefore, we might anticipate that a high number of ergative arguments would also be marked as topical, and in Katío this is the case: *ra* frequently occurs with ergative marker *ba* for a combination of *bara*.

As an example of this pattern, in (108a) the referent "two women" is introduced as an absolutive, and in (108b) it is in ergative case and is marked with *ra* as topical.

(108) a. *Maũne wũẽrã-rã ũme de ed'a k'o pan-a-si-a ãzha*
then woman-PL two house LOC stay AUX.PL-NPRS-PST-DECL 3PL

k'areba-d'a =ita.
help-PL =PURP
Then two women stayed in the house to help them.

b. *Mami wũẽrã-rã =ba =ra k'areba auda ẽ-b'-a-si.*
however woman-PL ERG =TOP help be.able NEG-AUX-NPRS-PST
However the women weren't able to help them. [Fight:8-9]

An effect of marking topicality is that it signals to the listener that the focused, most salient part of the utterance is elsewhere. This is related to Harms' (1994:190) and Mortensen's (1999:145) descriptions of *ra* marking a "shift in attention" or "directing attention away from the constituent to which it is attached." The following examples show the interaction of *ra* with focused constituents.

In example (109), the interrogative pronoun *k'ai* 'who' is the focused constituent of "who is under the house?" and in (109b), the answering constituent *wũãwũã* 'the child' is in focus. In both cases, the focused constituents are marked with *ta*, as seen in chapter 3. The focused constituents are underlined in this discussion.

(109) a. *K'ai =**ta** chu-pe de edre =ra?*
who =FOC STV-Q house under =TOP
Who is under the house?

b. *Wũãwũã* =***ta*** *de* *edre* *chu-a.*
child =FOC house under STV-DECL
The child is under the house.

Because the notions of focus and topic are by definition opposed, a focused constituent cannot be topical. We therefore expect that *ta* and *ra* never co-occur, and this is indeed the case in Katío.[7] Example (110) shows that *ra* cannot occur on a focused constituent. In (110b), marking the focused constituent with *ra* results in an inappropriate response to the question in (110a).

(110) a. *K'ai* =*ta* *chu-pe* *de* *edre* =*ra?*
who =FOC STV-Q house under =TOP
Who is under the house?

b. *Wũãwũã* =*ra* *de* *edre* *chu-a.*
child =TOP house under STV-DECL
The child is under the house (unacceptable in response to "who is under the house", as in (110a)).

However, the form of (110b) is grammatical when the first constituent is not in focus. Compare examples (110b) and (111b) below. Because the preceding question in (111b) is "where is the rock?", the location *de edre* 'under the house' is in focus, and not the constituent *mõg'ara* 'rock'. Now *mõg'ara* may be marked as topical since it is part of the pragmatic context of the question.

(111) a. *Mog'ara* *sãma* *chu-pe?*
rock where STV-Q
Where is the rock?

b. *Mog'ara* =*ra* *de* *edre* *chu-a.*
rock =TOP house under STV-DECL
The rock is under the house.

Example (112) is an excerpt from the beginning of a folktale about the creation of water, and it illustrates the syntactic and pragmatic functions of *ra* discussed so far. In

[7] I encountered one possible exception in the text El Pegante (Schöttelndreyer 1977:178) in which a word appears *netatara*, which could potentially be understood as *neta*=*ta*=*ra*, 'thing=FOC=TOP'. However, the correct morpheme breaks are unknown as in the text the word is simply glossed "things", and it is likely a different construction that does not require the mixing of two opposite notions in one word.

(112a), the time clause is marked as topical by *ra*, and the comment introduces the participant *Ãk'õre* 'God', who is in focus, marked with *ta*. In (112b), *dai* 'we' is in ergative case and is topical, contextually given, and thus is marked with *ba* and *ra*. Since *ãk'õre*[8] 'God' was previously mentioned, it is now marked with *ra* as topical, while the focus is on *K'arag'abi*. In (112c), *ãk'õre* 'God' has already been introduced and so is topical and now appears as an ergative marked with *ba* and *ra*, and another time clause is marked with *ra*.

(112) a. *Naã drua o-si-da =d'e =**ra** Ãkõre =ta chu*
this land do-PST-COMPL =LOC =TOP God =FOC STV

b'-a-si-a.
AUX-NPRS-PST-DECL
When this land was made, there was God.

b. *Dai =ba =**ra** ãk'õre =**ra** K'arag'abi a-ba-d'a.*
1PL.EXCL =ERG =TOP God =TOP K'arag'abi say-HAB-PL
We call God Karagabi.

c. *Maã ãkõre =ba =**ra** [naã drua =d'e nã-m-a-na] =**ra***
this God =ERG =TOP this land =LOC PROG-AUX-NPRS-COMPL =TOP

bi-ia o amae-si jõma-rã =ita.
very.good make leave-PST all-PL =BEN
This God, when [he] was in this land, [he] made and left it very good for everyone.

d. *Ãk'õre =ba =**ra** Jẽsẽra =á dia-si bania =**ra;** Jẽsẽra =ba*
God =ERG =TOP ant =DAT give-PST water =TOP ant =ERG

*mawũã jara-si: "Naã bania =**ra** mã aba =ita", a-si.*
thus say-PST this water =TOP 1SG one =BEN say-PST
God gave the water to the ant; the ant said this: "This water is for me only," he said.

e. *"Mã-re", a-si.*
1SG-POSS say-PST
"Mine", he said.

f. *"Wabema-rã =ita ẽ naã =**ra**".*
other-PL =BEN NEG this =TOP
"It's not for others, this."[9] [V&G:89]

[8] The inconsistent capitalization of *ãk'ore* is original to the text.

Finally, in (112d) *ãk'õre* 'God' is once again marked with *ba* and *ra*. Since *bania* 'water' is accessible from the context of the land and from the fact that it plays a large part in this traditional tale, it is also marked with *ra* while the dative recipient of the water, *jẽsẽra* 'the ant', is in focus. The second reference to *jẽsẽra* 'ant' is in ergative case, but is not marked with *ra*, perhaps because doing so would direct more attention away from it than desired as the ant takes an ergative role for the first time. The ant's reference to *naũ bania* 'this water', again marked with *ra* as topical, could potentially be interpreted as the syntactic topic of its clause.

In (112f), the final example from the excerpt above, the constituent marked with *ra* occurs to the right of the verbal position, occupied by the negative morpheme *ẽ*. This leads us to the final part of the discussion of the distribution of *ra*: postposing.

4.4.3 Postposing

In Embera Katío, elements often appear to the right of the verb, by a process that I refer to as postposing, and these are frequently marked with *ra*. As discussed in section 1.5.3, the basic word order of Embera Katío is SOV. However, the word order is often changed in order to highlight the pragmatic status of a certain constituent. When a constituent is in focus and should appear in the pre-verbal focus position, other constituents will often be moved to the right of the verb to free up the focus position slot and contribute to highlighting the focus on the pre-verbal constituent. When that happens, the postposed constituent is typically marked with *ra*, which marks it as topical.

Examples (113) and (114) show the postposing of an ergative and a dative argument, respectively. In example (113), *ra* occurs on the ergative constituent *usa=ba* 'dog=ERG', the subject of the transitive verb 'to kill' which has been moved to the right of the verb to make way for the focused content question word *k'ai* 'who' to be immediately pre-verbal. In a similar case, example (114) shows *ra* on a dative case constituent, *dama torro=á* 'to the white snake', which has been moved to the right of the verb.

[9] Since the majority of the texts in this study were written, I am unsure of the amount of pause that typically separates the postposed elements from the rest of the clause and therefore am unable to comment about intonation and pauses.

(113) *K'ai* =*ta* *bea-si* *usa* =*ba* =***ra?***
who =FOC kill-PST dog =ERG =TOP
Who did the dog kill?

(114) *DaiZeze* =*ba* *k'ãrẽ* =*ta* *jara-si* *dama* *torro* =*á* =***ra?***
God =ERG what =FOC say-PST snake white =DAT =TOP
What did God say to the white snake?

Another example of postposing is in the first clause of (112d) above, where *bania*=*ra* 'water=TOP' is postposed to be verb-final, leaving the pre-verbal focus position available for the dative constituent *jẽsẽra*=*á* 'to the ant', which is the focus of the utterance. A comparison of (114) with the first clause of (112d), repeated as (115), reveals how the focus structure affects the marking of the dative and absolutive arguments. In each case, *ra* marks the postposed constituent and the focused constituent is pre-verbal.

(115) *Ãk'õre* =*ba* =*ra* *Jẽsẽra* =*á* *dia-si* *bania* =***ra;***
God =ERG =TOP ant =DAT give-PST water =TOP
God gave the water to the ant; [V&G:p89]

In the same way, a postpositional phrase may also be moved to the right of the verb and optionally marked as topical, as in (117). In these cases of elicited data, the constituent marked with *ra* is not necessarily a previously mentioned constituent in the discourse, but may be understood to be marked as topical to emphasize that it is not in focus, corresponding to Mortensen's notion of *ra* as a nonfocal marker. Therefore the difference between (116) and (117) is stylistic, depending on how much attention the speaker wants to direct away from the phrase *puwɯrɯ ed'a* 'to the town'.

(116) *K'ai* =*ta* *jũẽ-si* *puwɯrɯ* *ed'a* =*ra?*
who =FOC arrive-PST town LOC =TOP
Who arrived in the town?

(117) *K'ai* =*ta* *jũẽ-si* *puwɯrɯ* *ed'a?*
who =FOC arrive-PST town LOC
Who arrived in the town?

In the following example from a text about how to build a house, the fact that the houses are made with trees is the focus, so the instrumental element *bakuru*=*ba* 'tree=INS'

is what comes in the pre-verbal slot. The absolutive element, *de* 'house', which is highly topical in a text about building houses, is moved to the right of the verb and marked with *ra*.

(118) *Ẽbẽra jõmaã̃ =ba bakuru =ba o pan-a-ri de*
person everyone =ERG tree =INS make AUX.PL-NPRS-HAB.PRS house

=ra.
=TOP
All Embera make houses with trees. [V&G:p54]

In summary, the function of *ra* is to mark elements associated with the pragmatic topic, including syntactic topics, subordinate clauses of time and condition, and postposed constituents. Instead of them being two homophonous morphemes, a default absolutive case marker and a pivot marker directing attention away from its host, *ra* is better understood as a single topic marker. Constituents are frequently marked as topical when they are postposed to highlight the focus on another constituent.

Although it is outside the scope of this thesis to posit that Katío is a 'topic-prominent' language as in Li and Thompson's typology based on topic vs. sentence prominence (1976:460), it is worth drawing attention to the number of topic-prominent characteristics that are at least apparently true of Katío. In addition to the existence of a topic marker, SOV word order is usually a feature of topic-prominent languages, and generally, "they have no dummy subjects, passive constructions are marginal (if they exist at all), zero NP-anaphora are not syntactically restricted..." (Gundel 1988:221). Further study is needed to investigate of the prominence of topic in Katío sentence structure and discourse.

CHAPTER 5

PROSODIC WORD DOMAIN AND PHRASE-LEVEL FUNCTION

In previous research, the morphemes *ba*, *á*, *ta* and *ra* have sometimes been called affixes and sometimes enclitics. Harms describes the counterparts of the case markers in Epena Pedee (SE) as "postpositional clitics" that "occur on the last word of a noun phrase and also on the verb of relative clauses" (1994:65), and Rex calls *ba*, *ta* and *á* enclitics (1975:40). Mortensen, on the other hand, calls them all suffixes (1999:8-9).

In fact, these morphemes have features usually associated with affixes as well as features usually associated with clitics. A feature of both affixes and clitics is phonological dependence on the host word (Velupillai 2012:93), and the case markers and pragmatic status markers in Embera Katío do lean phonologically on their host words. They do not have independent stress nor can they be independent utterances. Their phonological dependence is demonstrated by their affectedness by nasal spreading from the preceding word and the fact that they affect the stress pattern of the preceding word.

The process of nasal spreading extends to the limits of the phonological word, as shown in section 1.5.2. It affects the case marker and pragmatic status markers also, which I take as indicating that they are part of the phonological word. When these markers occur in a nasalized environment, those that begin with a blocking consonant, such as *ba* (119a) and *ta* (119b), are realized with a transitional nasal as described in 1.5.2, and those that begin with a permeable consonant like *ra* or a vowel like *á* are completely nasalized, as in example (119c).

(119) a. <*imama=ba*> [imãmãᵐba] 'jaguar=ERG'

b. <*k'ãrẽ=ta*> [χãn̆ẽⁿta] 'what=FOC'

c. <*mɯ=ra*> [mɯ̃n̆ã] '1SG=TOP'

Another important fact about these morphemes is that they are part of the prosodic domain of the words that precede them. As shown in (120) with ergative *ba* and topic marker *ra*, and in (121) with dative *á* and *ra*, the presence of these morphemes causes the placement of stress in the prosodic word to be shifted. A trochaic foot is aligned with the right edge of the prosodic word, and the stress pattern shows that these morphemes are included in that domain. When two occur after the same word, such as in (120c) and (121c), it is the first of the two that receives the stress of the trochaic foot.

(120) a. <*wũẽrã*> [ˈwũ̯ẽn̆ã] 'woman'

b. <*wũẽrã=ba*> [wũ̯ẽˈn̆ãᵐba] 'woman=ERG'

c. <*wũẽrã=ba=ra*> [wũ̯ẽn̆ãˈᵐbaɾa] 'woman=ERG=TOP'

(121) a. <*dama torro*> [damã ˈtoro] 'white snake'

b. <*dama torro=á*> [damã toˈro.a] 'white snake=DAT'

c. <*dama torro=á=ra*> [damã toro.ˈaɾa] 'white snake=DAT=TOP'

The traditional notion of clitic is that it is a "little word" with no independent accent that "leans accentually on an adjacent word" (Anderson 2011:2003). The fact that the case and pragmatic status markers of Katío are part of the prosodic word, affecting and receiving stress, is a characteristic usually associated with affixes, and would at first seem like a case for calling them so, as Mortensen does (1999:47, 144-145).

However, there are cases in which so-called clitics are stressed in order to maintain the foot structure of the prosodic word, such as in Modern Greek (Anderson 2011:2003). Anderson states that "In Modern Greek, enclitics do not usually receive stress... But when two such enclitics are attached to the same host, a stress appears on the penultimate one... This is a consequence of a general rule of Modern Greek that builds a trochaic foot over

two otherwise unstressed syllables at the right edge of a word, [as in (122c) and (123b),] provided the result does not involve a stress clash [as it would in (122b)]."

(122) a. [ˈðose] 'give!'
b. [ˈðose=mu] 'give me!'
c. [ˌðose=ˈmu=to] 'give it to me!'

(123) a. [triaˈðafilo] 'rose'
b. [triaˌðafiˈlo=mu] 'my rose'

His examples, shown in (122) and (123), are similar to Katío in that the enclitic morpheme is part of the domain of the prosodic word to which the language's stress rules apply. Thus, under the right circumstances, a clitic may be stressed.

According to Anderson (1992), "the mere fact of bearing stress would not be sufficient to establish the prosodic self-sufficiency of a given special clitic." He cites another example of stress-bearing clitics: Klavans (1980) shows an example from a dialect of Italian, spoken in southern Italy in which "a principle prohibiting stress further back from the end of a word than the antepenult results in stress being assigned to non-stem syllables in certain groups containing clitics, such as *sposàre=sé=lla* "marry=self=her; to get married to her"" (1992:204).

These two cases demonstrate that clitics may count for the prosodic word and may themselves receive stress in order to conform to the foot structure of the prosodic word. However, unlike Italian, which does not allow stress to precede the antepenult, Katío appears to prohibit stress preceding the penult, resulting in examples like (120b) and (121b). If the parallel may be drawn between these Greek and Italian clitics that are subject to the general stress rules at the level of the prosodic word and the morphemes in Katío, then the fact that *ba*, *á*, *ta* and *ra* affect and receive stress does not necessarily exclude them from clitichood, and is not enough to solidify their status as affixes.

While the inclusion of the case and pragmatic status markers of Katío in the prosodic word favors an analysis of affixes (but does not necessarily exclude them from being clitics), their distribution is more reminiscent of clitics, occurring at the right edge of a

nominal or clausal element, regardless of what class of word may be present on the right edge.

As in (121), where *á* and *ra* occur on the right edge of the noun phrase *dama torro* 'white snake', attaching to the adjective, *ba* and *ta* also occur on the right edges of phrases, as in example (124), where *ba* occurs on the post-nominal quantifier, and in (125), where *ta* occurs on the post-nominal adjective.

(124) *De o-d'a-i =ra nara aku̶-d'a-i bara ẽũjã =ta [zhi kima ũme]*
house do-PL-IRR =TOP first look-PL-IRR OBLG land =FOC 3SG spouse two

=ba.
=ERG
As for making a house, first [the two spouses] must look at the land. [V&G:p56]

(125) *[Burru zake] =**ta** arkila eta-pe, jira eta-si-a.*
donkey small =FOC rent take-SEQ hang take-PST-DECL
He rented a small donkey and rode it. [EP:35]

In a curious exception, *ta* does not behave like the others in that, while it attaches to the post-nominal adjective of an absolutive phrase as in (125), it does not attach to a post-nominal quantifier. See (126), where *ta* divides the quantifier *ũme* 'two' from the noun in modifies. This is distinct from the distribution of ba, which occurs on a quantifier in (124).

(126) *Ẽbẽra =**ta** ũme chiwu zru̶ga ed'a wã-na-pe, chiwu =ra zru̶ga-d'a-pe,*
person =FOC two goat steal LOC go-PL-SEQ goat =TOP steal-PL-SEQ

de ed'a ete-si-d'a.
house LOC take-PST-PL
Two Embera went to steal a goat, stole the goat, and took it to the house. [FD:1]

Example (127) is a similar case, in which *ta* divides the quantifier *ũme* 'two' from the noun it modifies.

(127) *Wũẽrã-rã =**ta** ũme jũẽ-si-d'a puwu̶ru̶ ed'a =ra.*
woman-PL =FOC two arrive-PST-PL town LOC =TOP
Two women arrived in the town.

The expected form with *ta* on the quantifier is shown to be ungrammatical by example (128).

(128) * *[Wũẽra-rã ũme] =ta jũẽ-si-d'a puwɯrɯ ed'a.*
woman-PL two =FOC arrive-PST-PL town LOC
Two women arrived in the town.

This is similar to the floating quantifiers of Japanese, in which a quantifier and its modified noun are commonly split by a case marker, as in example (129), where the subject marker *ga* separates the quantifier *san-nin* 'three-CLF' from its modified noun *gakusei* 'student'.

(129) *Gakusei ga san-nin ki-ta.*
student SBJ three-CLF come-PST
Three students came. [Nishiguchi (2009:156)]

According to Nishiguchi, there is a difference in scope and definiteness between a split NP and a non-split NP. "In Japanese, the use of a non-split quantifier phrase presupposes a unique set of entities, and thus corresponds to a definite description. On the other hand, the referents of a postnominal quantifier are not presupposed so that split quantifiers correspond to indefinites" (2009:156). Whether the apparent split noun phrases in Embera Katío, such as in (126) and (127), may have similar implications for scope and definiteness is beyond the scope of this study but would be relevant for future research.

As further criteria, I refer to Zwicky and Pullum's well-known six tendencies of clitics versus affixes (1983:503-504). These are not meant to be defining characteristics, but rather diagnostic tests:

(130) a. Clitics can exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems.
b. Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups.
c. Morphophonological idiosyncrasies are more characteristic of affixed words than of clitic groups.
d. Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups.
e. Syntactic rules can affect affixed words, but cannot affect clitic groups.
f. Clitics can attach to material already containing clitics, but affixes cannot.

I examine these tendencies one by one to compare them with the behavior of the case markers and pragmatic status markers of Embera Katío. Overall, these morphemes do not perform like affixes according to the tendencies proposed by Zwicky and Pullum (1983:503-504).

First Tendency: Clitics can exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems.

In Katío, case markers and pragmatic status markers show a low degree of restriction as to the syntactic category to which they can attach, with *ra* having the least restriction and *ta* the highest. The host is selected depending on what word is rightmost in the noun phrase or dependent clause.

We have seen the most common occurrences of all these morphemes on nouns and pronouns all throughout this thesis. We have also seen the occurrences of *ba* with a quantifier in (124), *á* with an adjective in (121), and *ta* with an adjective (125), and *ta* and *ra* on verbs of complement clauses or dependent clauses in sections 3.4 and 4.3. In addition to these, there are further instances like in (131), where the benefactive *ita* occurs on the quantifier of a noun phrase, in (132), where the dative *á* occurs on the auxiliary verb of a relative clause, and in (133), where the *ba* of reason occurs on the negative that marks the edge of a dependent clause.

(131) *Jẽsẽra =ba mawũã jara-si: "Naĩ bania =ra mɯ aba =**ita**", a-si.*
ant =ERG thus say-PST this water =TOP 1SG one =BEN say-PST
The ant said, "This water is for me only." [V&G:p89]

(132) *Mawũã krĩcha b'-ɯ =d'e [wũẽrã DaiZeze bedea ĩjã b'-ɯ*
do.thus think AUX-PRS =LOC woman God word believe AUX-PRS-DECL

*=**á** bed'ea-si-a DaiZeze =ta k'ãĩmokara =d'e.*
=DAT speak-PST-DECL God =FOC dream =LOC
As I am thinking like this, God spoke in a dream to [a woman who believed his word]. [BBT:5]

(133) *Diamase* *k'ãĩ-ña-mapai* *sopua* *nɯ-m-e-si-a* *[aɯtre* *plata* *ne*
night sleep-FUT-before sad PROG-AUX-INCP-PST-DECL more money thing

ẽ-ã] =***ba.***
NEG-DECL =REAS
Before I went to sleep that night I started feeling sad because [there was no more money]. [BBT:98]

This variety of syntactic categories to which the morphemes can attach is inconsistent with the behavior of affixes. The morphemes are functioning at a phrase-level rather than a word-level.

Second Tendency: Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups.

In Embera Katío there are no known arbitrary gaps in which a certain word may not host *ba*, *á* or *ra*. As shown in section 3.3, *ta* cannot occur on an element that is not an absolutive. As shown above in (128), it cannot occur on a postnominal quantifier either, at least in the cases shown. Because of these limitations, *ta* has the most affix-like profile of the group of morphemes. However, it can still be considered to be functioning at the phrase level because it occurs on postnominal adjectives and on complement clauses.

Third tendency: Morphophonological idiosyncrasies are more characteristic of affixed words than of clitic groups.

There is an allomorph *a* of the ergative case marker *ba* that always occurs on selected personal pronouns, discussed in chapter 2, but there are no other cases of a host + marker combination resulting in an unexpected morphophonological form.

Fourth Tendency: Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups.

There are no known cases of semantic idiosyncrasies where the meaning of the host + marker combination is other than the expected combination of the two independent meanings.

Fifth Tendency: Syntactic rules can affect affixed words, but cannot affect clitic groups.

I found no syntactic rule that would adequately test this tendency.

Sixth Tendency: Clitics can attach to material already containing clitics, but affixes cannot.

The pragmatic status marker *ra* can attach to material already containing case markers (e.g. examples (102) (113) and (114)). If the case markers are considered clitics, then *ra* cannot be considered an affix. However, there is no motivation for separating the case markers and *ra* into different syntactic categories as their behavior is parallel. If all are clitics, then this tendency presents no difficulty.

For Zwicky and Pullum's six tendencies for testing whether a phonologically bound morpheme is an affix or a clitic, the case markers and pragmatic status markers are overall less consistent with the behavior of affixes than of clitics.

Table 5 summarizes the behavior of the case and pragmatic status markers examined in this study. None are prosodically independent or may be independent utterances. All affect the stress of the preceding word and are affected by the process of nasal spreading. Syntactically, all occur on the right edge of a phrase, although *ta* is more restricted than the rest.

Table 5. Behavior of the case and pragmatic status markers of Embera Katío

	ba	***á***	***ita***	***ta***	***ra***
Prosodic independence	no	no	no	no	no
Independent utterances	no	no	no	no	no
Affects stress	yes	yes	yes	yes	yes
In domain of nasal spreading	yes	yes	yes	yes	yes
Occurs on phrase	yes	yes	yes	yes[1]	yes

Perhaps the most apt term for the mix of affix and clitic characteristics that we find in these Katío markers is the term phrasal affix or Juke's "affixal clitic." They behave strikingly similarly to the category described by Jukes (2006) (cited in Haspelmath (2015:11)) in Makassarese, which he terms "affixal clitics."[2] Jukes distinguishes these "affixal clitics" from the categories "affixes" and "clitics" because of their unique mix of characteristics: they count for stress like Makassarese affixes (example (134) but attach to phrases rather than words like Makassarese clitics (example (135)). The following examples are from

[1] With the noted exception of a phrase-final quantifier, as in (128).

[2] Haspelmath prefers not to use the terms "phrasal affix" or "affixal clitic" because they imply that the morphemes in question are a sub-type of affix or a sub-type of clitic respectively, and that is not true of the Juke's analysis, which places them in a third category entirely. Instead, Haspelmath coins the term "Afficlitic" for clarity in his discussion of Juke's third category (2015:11).

Jukes (2006:154-155). The accent mark on a vowel indicates stress, and the affixal clitics are written with boundary marker ≡.

(134) a. *bálla'*
house

b. *ballá'≡na*
house≡3.POSS
'his/her/their house'

(135) a. *miong le'léng≡ku*
cat black≡1.POSS
'my black cat'

b. *kalimbu' ta=tas-sungké≡ku*
mosquito.net NEG=NONVOL-OPEN≡1.POSS
'my mosquito net which is unopened'

However, as Haspelmath notes, while this third-category "affixal clitics" is useful for describing the unique categories of Makassarese, it will not solve our "classificatory problems" in general. In fact, even these three categories of affix, clitic and affixal clitic fail to describe all the idiosyncrasies of Makassarese (2015:11).

Although a precise syntactic categorization of these morphemes has not been given in this thesis, at least it is clear that they lean phonologically on their host words, forming part of the prosodic word for nasal spreading and stress patterns, and yet they function on a phrasal level. For the purposes of a language-specific description of Katío, we may provisionally call them affixal clitics, to borrow the term from Jukes (2006), in order to distinguish them from affixes.

CHAPTER 6

CONCLUSION

6.1 Summary

This study has summarized the descriptions of the morphemes *ba*, *ta* and *ra* in previous research in Embera Katío and related languages. Through analysis of published and collected texts and elicited material, I have tried to unify the previous analyses into succinct descriptions of the individual function of each morpheme.

While *ta* and *ra* are both considered to be the default absolutive case marker in Rex (1975) and Mortensen (1999) respectively, I have argued for an unmarked absolutive case in Katío and that *ta* and *ra* carry pragmatic status information only.

In chapter 2, I have shown that *ba* has three distinct functions: marking ergative, instrumental and reason. While there is semantic motivation for describing all three functions, at least historically, as one single *ba* morpheme, the terms associated with that semantic motivation, such as "agentive" or "ablative," fail to encompass all the applications of the three functions. This is especially evident in the instances where the ergative case marks arguments that are not semantic agents. Therefore, ergative case is a grammatical relation in Katío and is not purely semantically motivated by the notion of the agent.

While the absolutive distribution of *ta* makes it appear to be an absolutive case marker, I have shown that it only marks absolutives that are in focus, which is consistent with Mortensen's (1999) description. Embera Katío marks focus with pre-verbal default focus position, and *ta* has an analogous function as a focus marker and often redundantly marks focus on an absolutive in the default focus position. Marked focus for stronger emphasis is expressed with the morphemes *tru* and *buru*.

The function of *ra* is to code elements pertaining to the pragmatic topic. It marks the syntactic topic of a topic-comment sentence, and also frequently marks conditional clauses, adverbial clauses of time and constituents that have been postposed.

Finally, I have shown that the morphemes *ba*, *á*, *ita*, *ra* and *ta* have both the affix-like quality of being included in the prosodic domain of the preceding word in terms of nasal spreading and stress, and the clitic-like quality of functioning at a phrase level. Thus they are more adequately (but tentatively pending further research) described as affixal clitics that mark case and pragmatic status.

6.2 Implications for typological studies

The implications of this research for cross-linguistic typological studies include Embera Katío's classification in the World Atlas of Language Structure (WALS) (Dryer & Haspelmath 2013). Katío has not been included in the database yet, but two related languages have: Northern Embera Proper (called Emberá (Northern) in the WALS database), and Epena Pedee (SE).

For the feature 94A, Order of Adverbial Subordinator and Clause (Dryer 2013a), Emberá (Northern) is listed under "Suffixal adverbial subordinators," citing Loewen (1958:71, 76). However, Harms (1994:162, 161) is cited in the feature's article, with Epena Pedee (SE) shown as an example of a "language in which more than one of the above types occurs." A study of the many different clause subordinators is necessary for any conclusive decision on Embera Katío's classification. However, based on the fact that some adverbial clauses are subordinated with *ba*, *ra, ita* and *d'e*, which are not straightforwardly suffixes, I suspect that a classification like Epena Pedee's of multiple types of subordinators may prove to be more accurate for Northern Embera Proper and Embera Katío as well.

In Dryer (2013b), under the feature 51A, Position of Case Affixes, Emberá (Northern) is listed as having case suffixes, citing Loewen (1958:97-99). As discussed in chapter 5, the term suffix fails to adequately describe these morphemes. I believe that, given the limited classification options offered by WALS, "postpositional clitics" would be the most adequate available classification, given their phrasal distribution. However, Haspelmath points out that a large-scale cross-linguistic classification based on canonical definitions

will hardly be useful if "the great majority of elements will be noncanonical both as affixes and as clitics," just as we find in Katío (2015:17).

The behavior of Embera Katío's case and pragmatic status markers may prove relevant to the general discussion of the classification of the range of affixes and clitics and those morphemes that have features of both. I echo Haspelmath's (2015:17) question of whether the traditional notion and general definition of clitic is useful to us in light of the fact that it defines no theoretical category but is, as Zwicky (1994:xiii) calls it, an "umbrella term... for "problems", for phenomena that present "mixed" properties of some kind, not names for theoretical constructs." The notion of clitic does not appear to be useful for describing Katío except as an ad-hoc term to describe language-specific phenomena that are distinct from affixes. The behavior of these morphemes in Embera Katío, then, exemplifies the inadequacy of the traditional category of clitic as a cross-linguistic concept. This thesis contributes one more example of a language whose clitic-like phenomena are only partially described by the notion of clitic.

6.3 Areas for Further Research

This thesis has raised questions that should be pursued in future research.

The first is the syntactic behavior of *ba*, *ta*, *ra* and other locative morphemes such as *d'e* and *ed'a*, and how their distributions differ or correspond. This question was addressed briefly in chapter 5, but that discussion is by no means conclusive. It pertains to a discussion that is not currently clear-cut: the definition of the categories of affixes, clitics and everything in between. It will be important to the discussion of how to classify *ba*, *ta* and *ra* etc., to look at how they interact with coordinated noun phrases. Do they occur with each of the coordinated constituents or only once on the coordinated phrase? Unfortunately, for this study I did not have the data to explore the potentially revealing distribution of *ba*, *ta* and *ra* on coordinated phrases.

Another question for further research is how grammaticalized are *ta* and *ra*, since in many cases they appear to be optional. It would be valuable to test how the use and frequency of *ta* and *ra* vary in the speech of different speakers and regions. Are they quite regular in their occurrences even between different regions and generations, or is there

a high degree of variation between speakers and dialects? Does their use and frequency differ in different types of texts? Does competence in the use of *ra* indicate a higher level of "discourse competence" as Weber speculates for *-qa* in Huallaga Quechua (1989:402-403)?

A third area for future study is to investigate whether those locative morphemes that contain the segment *ba* might be analyzed as multi-morphemic, including *ba* with a sense of source, as suggested by Rex (1975:38). This may be the motivation behind Mortensen's term "ablative."

There are two locative morphemes with a sense of source that include the segment *ba*. The first is *ʉba* 'from', as shown in example (136). The question is whether or not *ʉ* and *ba* may be considered separate morphemes, and this is another occurrence of the same *ba* morpheme described in 2.

(136) *Zhi erbari bara-b'-e-si-bida, k'ore =ta wi ũme: k'ore =ta*
3SG pull PL-AUX-INCP-PST-EVID crocodile =FOC bear with crocodile =FOC

*do ed'a **ʉba**, wi =ta drua **ʉba**.*
river LOC from bear =FOC land from
He began to pull, it is said, the crocodile with the bear: the crocodile from in the river, and the bear from land. [V&G:p131]

A second locative *d'eba*, also meaning 'from' is potentially a combination of locative *d'e* and *ba*. The morpheme *d'eba* is nasalized to *neba* in example (137).[1]

(137) *Mawũã wã-i =ta krĩcha-pe ed'a tepadewa-si-a mʉ =ra ũtʉ ʉba*
do.thus go-IRR =FOC think-SEQ in start.walk-PST-DECL 1SG =TOP up from

k'atuma neba.
mountain from
After thinking like this about going, I started walking down from up on a hill. [BBT:47]

Exploring the difference in usage between these two locatives of source and defining the functions of all the locative and directional morphemes such as *d'e*, *ed'a*, *ʉba*, *id'u*, *edre*, etc. is also an important area for further research. A study of each of these locatives

[1] In this example, Mortensen wrote what I have separated as *ũtʉ ʉba* as *ũtʉ-pa*, 'up-ABL', analyzing it as *ba* of source. Since in Katío, adjacent identical vowels are combined (such as in the case of *de ed'a* 'house LOC', which is often written *ded'a*), I believe analyzing *ũtʉba* as a shortening of *ũtʉ ʉba* is quite plausible.

and their occurrences on noun phrases and dependent clauses may reveal keys for making generalizations about the system of Embera Katío and how these morphemes, like the case markers and pragmatic status marker, express relations not only between noun phrases but also between clauses.

The fact that these markers have features of affixes and clitics, as discussed in chapter 5, raises the question of how they should best be represented in the orthography. It would be valuable to design a study that tests whether writing them as affixes (as is currently the convention) or writing them as separate words, allows for greater reading fluency.

The morphemes *ba*, *ta* and *ra* are referred to as case markers and pragmatic status markers in this thesis. However, syntactically, they could be considered postpositions, just as researchers have referred to Japanese case markers as postpositions. This classification could complicate the formalism by positing that most noun phrases are postpositional phrases and thus the case marker itself is the phrase head, but further study may reveal that they are syntactically similar to locative postpositions in the language and that postposition is the best syntactic category for them.

In this thesis, I have not differentiated between the occurrences of the case markers and pragmatic status markers on clauses and their occurrences on noun phrases. Since it is common for adpositions to become complementizers of clauses, and since *ba*, *ta* and *ra* can occur on clauses, they may have become or may in the process of becoming complementizers. It would be valuable to do a study of the clause-combining strategies of Katío and see how the clausal occurrences of *ba*, *ta* and *ra* compare to other subordinating strategies. If the occurrences of *ba*, *ta* and *ra* on clauses correlate more with complementizing strategies than with their noun phrase occurrences, we might consider them to be grammaticalized.

REFERENCES

Adelaar, Willem F.H. 2004. *The languages of the Andes.* Cambridge: Cambridge University Press.

Aguirre Licht, Daniel. 2009. Choco languages. In Keith Brown & Sarah Ogilvie, (eds). *Concise encyclopedia of languages of the world,* 225-38. Oxford: Elsevier.

Aguirre Licht, Daniel. 1999. Ergatividad, focalización, tematización y topicalización en Embera. *Congreso de Lingüística Amerindia y Criolla* 3. 317-30. Bogotá: Universidad de Los Andes.

Anderson, Stephen R. 1992. *A-morphous morphology.* (Cambridge studies in linguistics.) New York: Cambridge University Press.

Anderson, Stephen R. 2011. Clitics. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice, *The Blackwell Companion to Phonology,* 2002-18. Malden, MA: Wiley-Blackwell. http://cowgill.ling.yale.edu/sra/clitics_ms.pdf. (6 July, 2015.)

Butt, Miriam. 2006. The dative-ergative connection. In Bonami, O & P. Cabredo-Hofherr (eds), *Proceedings of the Colloque Syntax-Semantique Paris (CSSP).* 6:69-92. The Hague: Thesus.

Campbell, Lyle. 1997. *American Indian languages: the historical linguistics of native America.* (Oxford studies in anthropological linguistics.) New York: Oxford University Press.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li, *Subject and topic,* 25-55. New York: Academic Press.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Cohn, Abigail. 1993. The status of nasalized continuants. In Huffman, M. & Krakow, R., *Nasals, nasalization and the velum,* 329-67. San Diego: Academic Press.

Costenla Umaña, Adolfo & Enrique Margery Peña. 1991. Elementos de fonología comparada Choco. *Filología y lingüística* 17. 137-91.

Crystal, David. 1997. *A dictionary of lingusitics and phonetics.* 4 edn. Oxford: Blackwell Publishers.

Dixon, Robert M. W. 1994. *Ergativity.* Cambridge: Cambridge University Press.

Dixon, Robert M. W. 2011. *The languages of Australia.* Cambridge: Cambridge University Press.

Dooley, Robert A. & Stephen H. Levinsohn. 2001. *Analyzing discourse: a manual of basic concepts.* Dallas, TX: SIL International.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547-619.

Dryer, Matthew S. & Martin Haspelmath, Eds. 2013. *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info. (6 July, 2015.)

Dryer, Matthew S. 2013a. Order of adverbial subordinator and clause. In Matthew S. Dryer & Martin Haspelmath, (Eds). *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info. (6 July, 2015.) http://wals.info/chapter/94. (20 June, 2015.)

Dryer, Matthew S. 2013b. Position of case affixes. In Matthew S. Dryer & Martin Haspelmath, (Eds). *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info. (6 July, 2015.) http://wals.info/chapter/51. (20 June, 2015.)

DuBois, John W. 1987. Discourse basis of ergativity. *Language* 63. 805-55.

Givón, Talmy. 1984. *Syntax: a functional-typological introduction.* 1. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Greenfield Velez, Gisella T. 2012. *The acoustic qualities of Embera Katío stops.* Grand Forks, ND: University of North Dakota thesis.

Grimm, Scott M. 2005. *The lattice of case and agentivity.* Amsterdam: Universiteit van Amsterdam thesis.

Gundel, Jeanette K. 1988. Universals of topic-comment structure. In Michael Hammond, Edith A. Moravcsik & Jessica R. Wirth, *Studies in syntactic typology,* 209-35. Amsterdam: John Benjamins.

Gundel, Jeanette K. & Thorstein Fretheim. 2004. Topic and focus. In Laurence Horn & Gregory Ward, *The Handbook of Pragmatics,* 175-96. Oxford: Blackwell.

Gunn, Robert D., Ed. 1980. *Clasificación de los idiomas indígenas de Panamá, con un vocabulario comparativo de los mismos.* 7. (Lenguas de Panamá.) Panama: Instituto Nacional de Cultura, Instituto Lingüístico de Verano.

Haiman, John. 1978. Conditionals are topics. *Language* 54(3). 564-89.

Harms, Phillip. 1994. *Epena Pedee syntax: studies in the languages of Colombia 4.* Arlington, Texas: Summer Institute of Linguistics and University of Texas at Arlington.

Harris, Alice C. 2002. *Endoclitics and the origins of Udi morphosyntax.* Oxford: Oxford University Press.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86. 663-87.

Haspelmath, Martin. 2015. Defining vs. diagnosing linguistic categories: a case study of clitic phenomena. In Joanna Błaszczak, Dorota Klimek-Jankowska & Kryzsztof Migdalski, *How categorical are categories?,* np-np. Berlin: De Gruyter Mouton.

Huber, Randall Q. & Robert B. Reed. 1992. *Comparative vocabulary: selected words in indigenous languages of Colombia.* Santafé de Bogotá, Colombia: Instituto Lingüístico de Verano.

Jukes, Anthony. 2006. *Makassarese (basa Mangkasara'): a description of an Austronesian language of South Sulawesi.* Melbourne: University of Melbourne dissertation.

Kim, Alan H. O. 1988. Preverbal focusing and type XXIII languages. In M. Hammond, E. Moravcsik & J. Wirth, *Studies in syntactic typology,* 147-72. Amsterdam/Philadelphia: John Benjamins.

Klavans, Judith L. 1980. *Some problems in a theory of clitics.* Bloomington: Indiana University Linguistics Club.

Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus, and the mental representation of discourse referents.* (Cambridge Studies in Linguistics 71.) Cambridge: Cambridge University Press.

Lewis, M. Paul, Gary F. Simons, Charles D. Fennig, Eds. 2014. *Ethnologue: languages of the world.* 17th edition edn. Dallas, Texas: SIL International. http://www.ethnologue.com. (6 July, 2015.)

Li, Charles N. & Sandra A. Thompson. 1976. Subject and topic: a new typology of language. In Charles N. Li, *Subject and topic,* 457-89. New York: Academic Press.

Loewen, Jacob A. 1963a. Choco II: phonological problems. *International Journal of American Linguistics* 29(4). 357-71.

Loewen, Jacob A. 1963b. Chocó I: introduction and bibliography. *International Journal of American Linguistics* 29(3). 239-63.

Loewen, Jacob A. 1958. *An introduction to Ẹpẹra speech: Sambu dialect.* Seattle: University of Washington dissertation.

Mejía Fonnegra, Gustavo. 2000. Lenguas de la costa pacífica. In Maria Stella González de Pérez & María Luisa Rodríguez de Montes, (González de Pérez, María Stella and María Luisa Rodríguez de Montes). *Lenguas indígenas de Colombia.* Santafé de Bogotá: Instituto Caro y Cuervo.

Morris, Judith Kay. 1977. *Nasality in Ẽperã.* Dallas, TX: University of Texas at Arlington thesis.

Mortensen, Charles A. 1999. *A reference grammar of Northern Embera languages.* Arlington, TX: SIL International and University of Texas at Arlington.

Nishiguchi, Sumiyo. 2009. Quantifiers in Japanese. In Peter Bosch, David Gabelaia & Jérôme Lang, *Logic, language, and computation,* 153-64. Heidelberg: Springer-Verlag Berlin.

Pardo Rojas, Mauricio & Daniel Aguirre Licht. 1993. Dialectología Chocó. Estado actual de la clasificación de las lenguas indígenas de Colombia. Bogotá, Colombia: Instituto Caro y Cuervo.

Payne, Thomas E. 1997. *Describing morphosyntax.* Cambridge, United Kingdom: Cambridge University Press.

Reesink, Gerard P. 1981. Review article: the Whorfian hypothesis and Siroi grammar: a review of: Siroi grammar, by M.A. Wells. *Language and Linguistics of Melanesia* 13. 81-105.

Rex, Eileen Elizabeth. 1975. *On Catio grammar.* Arlington, TX: University of Texas at Arlington thesis.

Rex, Eileen and Mareike Schöttelndreyer. 1972. Sistema fonológico del Catío. In Florence L. Gerdel et al.*Sistemas fonológicos de idiomas colombianos 2,* Lomalinda, Colombia: Editorial Townsend.

Rivet, Paul. 1943. La lengua Choco. *Revista del Instituto Etnológico* 1. 13-96, 297-349.

Sanchez A., Olga Castro G. & Reinaldo Binder. 1977. *Una gramática pedagógica del Waunana.* (Lenguas de Panama 3.) Panama: Instituto Nacional de Cultura, Dirección del Patrimonio Histórico; Instituto Lingüístico de Verano.

Schöttelndreyer, Mareike. 1977. translated by Ruth Monterroso A. *La narracion foklorica catia como un drama en actos y escenas.* Loma Linda, Colombia: Editorial Townsend.

Schöttelndreyer, Marieke & Stephen H. Levinsohn. 1976. The Catío folktale as a play in acts and scenes. *Poetics* 19. 247-80.

Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax: structure, meaning and function.* Cambridge: Cambridge University Press.

Velupillai, Viveka. 2012. *An Introduction to linguistic typology.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Villalobos, Rito Llerena and Ernesto Llerena Garcia. 2004. *Ẽbẽra bedea zõrarã ena bema: diccionario etnolingüístico de la lengua Ẽbẽra del Alto Sinú.* Montería, Colombia: Convenio Interadministrativo URRA S.A.E.S.P.

Weber, David J. 1989. *A grammar of Huallaga (Huanuco) Quechua.* (University of California Publications in Linguistics 112.) Berkeley: University of California Press.

Zwicky, Arnold M. 1994. What is a clitic? In Joel A. Nevis, Brian D. Joseph, Dieter Wanner & Arnold M. Zwicky, *Clitics: a comprehensive bibliography 1892-1991,* xii-xx. Amsterdam: Benjamins.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English N'T. *Language* 59(3). 502-13.